

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: June 23, 2003, 12:51:58 ; Search time 8181 Seconds
(without alignments)
11586.339 Million cell updates/sec

Title: US-09-817-198b-1
Perfect score: 3257
Sequence: 1 tgcgcgtgccccgcgcag.....aaaaaaaaaaaaaaaaaa 3257

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_nam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2696.6	82.8	190517	9	CNS01DX4	AL139022 Human chr
2	1192.8	36.6	7924	6	AX347040	AX347040 Sequence
3	1192.8	36.6	7924	6	AX348456	AX348456 Sequence
c	1160.8	35.6	7924	6	AX347041	AX347041 Sequence
c	1160.8	35.6	7924	6	AX348457	AX348457 Sequence
6	1000.2	30.7	1054	6	AX399903	AX399903 Sequence
7	866.2	26.6	895	6	AX014147	AX014147 Sequence
8	813	25.0	3139	10	BC027769	BC027769 Mus muscu
9	577.6	17.7	945	10	RATRA15X	M83679 Sprague-Daw
10	477.4	14.7	481	6	AX396088	AX396088 Sequence
c	438.8	13.5	463	6	AX334820	AX334820 Sequence
c	423.6	13.0	162495	2	AC096084	AC096084 Rattus no
13	416.4	12.8	2560	10	BC013790	BC013790 Mus muscu
c	378.4	11.6	257226	2	AC124556	AC124556 Mus muscu
15	307	9.4	313	6	AX341683	AX341683 Sequence
16	234.2	7.2	2048	9	EC002977	EC002977 Homo sapi
17	234.2	7.2	2818	9	AK025165	AK025165 Homo sapi
18	230.6	7.1	1980	9	S53268	S53268 Homo sapien
19	229	7.0	624	9	AF498943	AF498943 Homo sapi
20	229	7.0	660	9	HSRAB8	X56741 H.sapiens m
21	227.4	7.0	638	10	S53270	S53270 MEL-RAS-rel
22	224.2	6.9	1337	10	BC019990	BC019990 Mus muscu
23	222.6	6.8	760	4	CFRAB8	X56385 Canine rab8
24	218.2	6.7	765	5	DYGORA2	M83677 Sprague-Daw
25	217.8	6.7	624	6	AX236078	AX236078 Sequence
26	217.8	6.7	1128	9	BC020654	BC020654 Homo sapi
27	217.8	6.7	1161	6	AX236076	AX236076 Sequence
28	217.8	6.7	1265	9	AB038995	AB038995 Homo sapi
29	217.8	6.7	2497	6	AX285074	AX285074 Sequence
30	217.8	6.7	2497	6	AX285080	AX285080 Sequence
31	217.8	6.7	2497	6	AX285089	AX285089 Sequence
32	217.8	6.7	2497	6	AX285091	AX285091 Sequence
33	217.8	6.7	3077	9	AK001111	AK001111 Homo sapi
34	216.2	6.6	3110	9	HSR04678	AL833365 Homo sapi
35	211.4	6.5	740	10	RN053475	U53475 Rattus norv
36	203.2	6.2	991	10	RATRA10X	M83677 Sprague-Daw
37	201.8	6.2	897	10	AF035646	AF035646 Mus muscu
c	198.6	6.1	110000	2	AC048347_0	AC048347 Homo sapi
c	198.6	6.1	110000	2	AC048347_1	Continuation (2 of
40	198.6	6.1	302718	2	AC027292	AC027292 Homo sapi
41	198.4	6.1	1029	9	BC000896	BC000896 Homo sapi
42	198.4	6.1	3533	9	AK023223	AK023223 Homo sapi
43	196.8	6.0	3521	9	HSR01620	AL136650 Homo sapi
44	196.6	6.0	2048	3	AY069671	AY069671 Drosophila
45	196.6	6.0	2210	3	D84347	D84347 Drosophila

ALIGNMENTS

RESULT 1
CNS01DX4
LOCUS
DEFINITION
Human chromosome 14 DNA sequence BAC R-84019 of library RPCI-11
from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION
AL139022
VERSION
AL139022.4 GI:14148954
KEYWORDS
HTG.
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 190517)
AUTHORS
Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
Brottier,P., Catolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,

Levy, M., Eckenberg, R., Bruls, T., Deberardinis, V., Cruaud, C., Gysapay, G., Saurin, H. and Weissenbach, J.
 Sequencing of the human chromosome 14
 Unpublished
 2 (bases 1 to 190517)
 Genoscope.
 Direct Submission
 Submitted (17-MAY-2001) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 On May 18, 2001 this sequence version replaced gi:10798506.
 ----- Genome Center
 Center: Genoscope / Centre National de Sequencage
 Center code: GS
 Web site: http://www.genoscope.cns.fr/
 Contact: Segref@genoscope.cns.fr

 The following BAC sequence is oriented from the T7 to the SP6 end.
 Upstream BAC (overlapping the T7 end) : C-2509G16 (AC-AL355076)
 Downstream BAC (overlapping the SP6 end) : C-2300N4 -----
 Summary Statistics
 Assembly program: Phrap; version 2.0
 Quality coverage: 7.14x in Q20 bases; sum-of-contigs

 Overall quality chart :
 Range : bases
 0 :
 1 : 9
 10 : 19
 20 : 29 : 45
 30 : 39 : 444
 40 : 49 : 4501
 50 : 59 : 11521
 60 : 69 : 12036
 70 : 79 : 23986
 80 : 89 : 56256
 90 : 99 : 81728

Percentage of bases with a quality value >= 40 : 99 %.

FEATURES source

1. 190517
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="14"
 /clone="R-840119"
 /clone_lib="RPCI-11"
 22163. .22307
 /note="matching EMBL:H99639
 RHdb:RH68943
 dbSTS:STS48849
 Identified using the e-PCR software (G. Schuler)"
 30401. .30624
 /note="matching EMBL:R42375
 RHdb:RH69913
 dbSTS:STS49790
 Identified using the e-PCR software (G. Schuler)"
 30772. .30949
 /note="matching EMBL:AA053315
 RHdb:RH67850
 dbSTS:STS47777
 Identified using the e-PCR software (G. Schuler)"
 43552. .43739
 /note="matching EMBL:Z39082
 RHdb:RH25206
 dbSTS:STS25011
 Identified using the e-PCR software (G. Schuler)"
 44712. .44852
 /note="matching EMBL:L00635
 RHdb:RH53870
 dbSTS:STS42913
 Identified using the e-PCR software (G. Schuler)"
 44712. .44832

Query Match 82.8%; Score 2696.6; DB 9; Length 190517;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 2717; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 473 TGGCATGGACTTCTATGAACAAGTGCCTGCACCAACCTCAACATTAAGAGTCAATTCAC 532
 DB 157619 TCCAGGGGAGAGGAGAGACACTGGACTACCTGTGCCCTTGGTTCCAGTCAATTCAC 157678

QY 533 CGGTCTGACAGAGCTGGTGTGCGAGGCCCATAGGAAGAGCTGGAAGCCCTCCGGATCGG 592
 DB 157679 CGGTCTGACAGAGCTGGTGTGCGAGGCCCATAGGAAGAGCTGGAAGCCCTCCGGATCGG 157738

QY 593 TGCCAGCAATGAGTTGGCACTGGCAGAGCTGGAGGAGGAGGCGCAACCCGAGGGGCC 652
 DB 157739 TGCCAGCAATGAGTTGGCACTGGCAGAGCTGGAGGAGGAGGCGCAACCCGAGGGGCC 157798

QY 653 AGCGAACTCTTCGAAACCTGCTGTGTGAGTCTGTGTGGGCGACCCACACGACACACC 712
 DB 157799 AGCGAACTCTTCGAAACCTGCTGTGTGAGTCTGTGTGGGCGACCCACACGACACACC 157858

QY 713 CCTCTTCCCTCAGGAGGCCCGTGGGAGAGGAGGCGGGGCTTTGCCCTGTGCTGTGT 772
 DB 157859 CCTCTTCCCTCAGGAGGCCCGTGGGAGAGGAGGCGGGGCTTTGCCCTGTGCTGTGT 157918

QY 773 CCTCTGTGTGATGACCTTATGAGTATCAGTACCTACTTCCCTGTGCTGTGTGT 832
 DB 157919 CCTCTGTGTGATGACCTTATGAGTATCAGTACCTACTTCCCTGTGCTGTGTGT 157978

QY 833 GAGCGGCTCTGCTGTCTATCTCAAGCAGCCCTGTGCCAGCCCTGCCAGCCCTGCCAGTGT 892
 DB 157979 GAGCGGCTCTGCTGTCTATCTCAAGCAGCCCTGTGCCAGCCCTGCCAGTGT 158038

BASE COUNT 52747 a 43175 c 42572 g 52023 t
 ORIGIN

/note="matching EMBL:L00635
 RHdb:RH53755
 dbSTS:STS12411
 Identified using the e-PCR software (G. Schuler)"
 47222. .47411
 /note="matching EMBL:AA165116
 RHdb:RH48540
 dbSTS:STS41599
 Identified using the e-PCR software (G. Schuler)"
 128322. .128437
 /note="matching EMBL:G07650
 RHdb:RH34434
 dbSTS:STS6376
 Identified using the e-PCR software (G. Schuler)"
 164135. .164257
 /note="matching EMBL:AA160869
 RHdb:RH48489
 dbSTS:STS41548
 Identified using the e-PCR software (G. Schuler)"
 169170. .169323
 /note="matching EMBL:W56308
 RHdb:RH47847
 dbSTS:STS40906
 Identified using the e-PCR software (G. Schuler)"
 171252. .171384
 /note="matching EMBL:Z39818
 RHdb:RH10964
 dbSTS:STS1017
 Identified using the e-PCR software (G. Schuler)"
 173564. .173713
 /note="matching EMBL:R10900
 RHdb:RH53901
 dbSTS:STS22552
 Identified using the e-PCR software (G. Schuler)"
 173883. .174012
 /note="matching EMBL:G27797
 RHdb:RH32719
 dbSTS:STS4675
 Identified using the e-PCR software (G. Schuler)"

Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nigri.nih.gov

Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granter, S., Guan, X., Gupta, J., Hagnighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 55 Row: d Column: 5
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

FEATURES

Location/Qualifiers

1..3139
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N"
 /clone="MGC:38375 IMAGE:5345297"
 /tissue_type="mammary tumor, C3(1)-Tag model. Infiltrating
 ductal carcinoma. 5 month old virgin mouse."
 /clone_lib="NCI_CGAP_Mam6"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 208..846

CDS

/codon_start=1
 /product="unknown (protein for MGC:38375)"
 /protein_id="AAH27769.1"
 /db_xref="gi:20380722"
 /translation="MAKQYDVLFRLLIGDSVGKTCILLCRTDNEFHSHSTIGVD
 FKMTIDVGIKVIQWDAQERQYITKQYRRAQGLFLVDISERSYQHIMKW
 VSDVDEAPGVKILIGNKADQKRGQKQOLAKYGMDFYETSACTNLNIKE
 SFRLRLVLAHAKELDLFRASNELALAELEDEKPGPSNSKTCWC"
 723 a 847 c 878 g 691 t

BASE COUNT

ORIGIN

Query Match 25.0%; Score 813; DB 10; Length 3139;
 Best Local Similarity 67.4%; Pred. No. 2,1e-198;
 Matches 1662; Conservative 0; Mismatches 640; Indels 165; Gaps 30;
 22 TCCCGGCCCGTGGCCCGATGCGGAGAGCAGTACGATGCTGTTCGGCGTGCCTC 81
 185 TCGGGCCCGCTTGAGTCCCTGCCATGCGGGAACAGTACGATGCTGTTCGGCTACTGC 244
 82 TGATCGGGGACTCCGGGTGGCAAGACCTGCTGCTGCGCTTACCGACACACGAGT 141
 245 TGATCGGGGACTCCGGGTGGCAAGACATGCTGCTGCGCTTACCGACACACGAGT 304
 142 TCCACTCTCGACATCTCCACCATCGGTGTGACTTTAAGATGAAGACCATAGAGGTAG 201
 305 TCCACTCTCGCATATCTCCACCATCGGTGTGACTTTAAGATGAAGACCATAGAGT 364
 202 AGGGCATCAAGTGGGATACAGATCTGGACATCGGACATCGAGGCGAGGATACAGACCA 261
 365 AGGGCATCAAGTGGGATACAGATTTGGGACACAGCAGGCGAGGAGGTACAGACTA 424
 262 TCACAAAGCAGTACTATCGGGGGCCAGGGGATATTTTGGTCTATGACATTAGCAGCG 321
 425 TCACAAAGCAGTACTATCGGGGAGCCAGGGGATATTTTGTCTACGCATTAGCAGTG 484
 322 AGCGCTCTTACGACATCATGAAGTGGGTGAGTGCAGTGCAGTGCAGTGCAGCAGAG 381
 485 AGCGCTCTTATCAGCATATCATGAAGTGGGTGAGTGCAGTGCAGTGCAGTGCAGAG 544
 382 GCGTCCAGAGATCTTATGGGAATAAGCTGATGAGGACGACAGAGAGGAGGTGGGAA 441
 545 GAGTCCAGAGATCTTATGGGAATAAGCTGATGAGGACGACAGAGAGGAGGTGGGGA 604
 442 GAGAGCAAGGGCAGCAGCTGGCGAAGGAGTATGGCATGGACTTCTATGAACAAGTGCCT 501

605 GAGAGCAGGGGCGACGAGCTGGCTAAGAGTACGGCATGGACTTCTACGAACAAGTGCCT 664
 502 GCACCAACCTCAACATTAAGAGTCAATTCACGGGTCTGACAGAGCTGCTGCAGGCCCC 561
 665 GCACCAACCTCAATTAAGAGTCAATTCACGGGTCTGACAGAGCTGCTGCAGGCCCC 724
 562 ATAGGAAGAGCTGGAAGGCTCCGGATGGTGGCAGCAATGAGTTGGCACTGGCAGAGC 621
 725 ACAGAAAGAGCTGGATGGTCTCCGAACACGTGCGACAGAGCTGCGACTGGCCGAGC 784
 622 TGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 681
 785 TGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 844
 682 GAGTCTGTGTGGGGCACCCACACAGCACCCCTCTTCCCTCAGAGAGGCGCTGGGCGAGA 741
 845 GAGGTCTGAGTGGAGTCCCCC---CAGGCCCTTCCCTCAGAGAGGCGCTGGAGGCA 900
 742 CAGGGAGCGGGGCTTTGGCCCTGCTGCTGCTCTCGTGTGATGAGACCTTATGAGTATC 801
 901 ---GGGAGTCTGGGCTTTGGCC-ACCACCGCTCTTCAATTTGACATCTCTGTTGAAAAC 956
 802 AGTAGCACACTACTCCCTGCTGCGCTGAGAGCGCTGCTGCTGCTCATCTC-AAAGCAGC 860
 957 AGTACTGCTACT-CCCTGCTGGCCCTGGAGCAGCCCTGCTGCTGCTGCTGCTGAGCAGC 1015
 861 CCTGTGCTCCAGCGCCGTC-CACCTGGAGTGGTCTTCTCAGCCTGTTTCCCGACGACCA 919
 1016 CTCTGCCCCCAGGCCCTTAATTTCTGGAGTGGTGACCTTTTCCAGCCTGTTTCCCGACGCA 1075
 920 GGCCTGTACAGACCCCGACGATGCGCGAAGCACTGTCTCACCACCTCCCGACCCACGAG 979
 1076 GGCCTCTGTGA-CCCCAGATGTGCTGC-----GACCCACCCAGGCTCTCTGGG 1123
 980 ACAACAGCAGGCTGGAGTCCAGCGCACCTTTCAGTCTCTCTTCTCCGTCGATCTGT 1039
 1124 ACGAGGTCAAGGTGGAGTCAAGCCACCTCAAG-----ATTCTTAGTGCATCTCGT 1177
 1040 CTCTCTCTGCTTTTCTCTTCTCCCGACCTTCTTCTCTGACCCCTCCCTCCCGGTG 1099
 1178 GGGCTCTGCTTTTCTCTCT-TCCTTCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCT 1233
 1100 GCTTCTGATCAAGCTCTCAACCCCGTCCCGCTGCTGCTCTCTCTCTCTCTCTCTCTCT 1159
 1234 TAATTCGACCCCTATCATCTCCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1293
 1160 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1219
 1294 TGCTACTGCACTTCTCT-TCCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1351
 1220 CTTGGATCCAGGAAGAACCTCCACCTGCTCTGCTGGTGGGCGCAAGGCTACAGGTG 1279
 1352 GCTCTGCCCCAGGAGGAGTCTTCACTCTG-CTCGGGGTGGGCTGAGGCTA-AGGCTG 1409
 1280 CTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1339
 1410 CTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1446
 1340 ACCTCGCAAGTGGAGCATGAGGTAGGAGGAAACAGCAACCGGGGAGTCTCTCAGGCT 1399
 1447 GCCTGGGAAGGTAGAGCACTGAGGTAGGAAGAAACACTA--CTAGGTGGCCCTCAAGCC 1504
 1400 GGGGTGCTCTACCTCTACCTATTCCTCCCGACAGAGCTTTGCCCTTGGCTGGCTGCCCGC 1459
 1505 TGGACTGGCCCTCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1564
 1460 CTGCTCTTTTGGGAACCTGAGCTCAGAGCGAGGTGCTTTCAGAGAGGAAACAAATGA-- 1517
 1565 GTGCTCTGTGAGGACTGAGTCCAGAGCGAGGACACAGCAGAGAGGAAACAGAGAA 1624
 1518 ---GGGTGGCAGGATAAAGTACCTCCATCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1573

```

Db      1625 CGATGGGTCGACGGGAT--TAAGTCACCTCTGTCTCTACCTCCCATGCCCTC----- 1675
Qy      1574 CAATTTCTCCACCTGGCTCCCAAAATTAAGATGTGGACCAAGGCCCTGTGGTACTCC 1633
Db      1676 ----TCTCCTCACCTGGATCACAATTTAAAGATGGGATCGAGGCTGAGGACAC-CA 1730
Qy      1634 AGGGCCAGGAGAGACCCCTGGGTCAGTCACACTGTGAGCCCAACCCATCGACTCCACAAG 1693
Db      1731 AAGGGCAAGGGAGCCCTGGGGTCA--GATACTGTATAGTCAGCTGTGAATCCAGATAG 1788
Qy      1694 GGGAGCATTTTGGAAATGAAGACTAGCTCTCTATCATCAGGTTAAGAGCAGGGAGAGCT 1753
Db      1789 AGGCAAGCGGATGCGAAGGCGCGCTCTAGAACAGGTTAGACGAGAGCAGAGGG 1848
Qy      1754 GGCCAGGACAGCAGTTTGCACAGCAGAGGGAATGTAGCAACAGCAGGGCCCTCTAGGC 1813
Db      1849 AGC-----TTGCGAGAACAGCAGCTGCTGTGGGACTACAGGCGCATCAGAGGT 1896
Qy      1814 CCCATCTTCCATTCTTAGGTTAAGAGAGCATTTCTCAGACTCCAGGGCGGAGACTGA 1873
Db      1897 GACAGTAATC-----AACCACCTCCTCTTTCACTGA 1927
Qy      1874 GCCTAGCTTTCAGCAACCAAGGTTCTCTGGGACCCAAAGTTTATGGGAGAGGCAAG 1933
Db      1928 ACCAGCCTTAGGAGTTAGGTTTCCCGAGCTCTGGACACAGGCCCTTAAGACAGGG 1987
Qy      1934 ACTTCATGGGAAGAGAGAGGAGGCCCTGGGTAGAAACGCTTGCTGCTCTCTTTGG 1993
Db      1988 AAGAGAGGGTGGGTA-----TGATGGAATACTTGGTGTCTTTCGCTTC 2034
Qy      1994 CTTTAAAGCAAAACGGCTCATCTTGGCTCTACCTCTGATAGCTTTGAGGGTTTGCAA 2053
Db      2035 CAG-----GACTCCCTCCCTGGTTCCTTATTTCTTAAGTCGGCTTTGCAA 2082
Qy      2054 CCACACTGTGCTACAGGTGAGGAGAGAGACTCTTCCCTCCAGAGTGTCTGTACAGG 2113
Db      2083 CTGCTGTGTGCTACCAAGGAGGAGGAGAGACACCTCTCCAGAGTCTATGCTCAGG 2142
Qy      2114 AAGTTCTTTAAACCCATATGCCCCAAGAGTAGCTGTGAGAGGCCCTTTAAAGACGAA 2173
Db      2143 AAGTTCTTTAAACCC--TGTGTTCAAGAGTAGCTTGAAGGAGGCCCTGAAACAAAC 2200
Qy      2174 CAAGTAATTTA-----CCAGTTCTACTGGGTTCTTCCCAACCTCCCAAGTGGCGAG 2228
Db      2201 AAACAACAAACAAAGCCACCCCTGGTAGGTTTCATCTTACATTCCTCAAGTGGGAGGA 2260
Qy      2229 GCCTAAGAGAGGCTCATTTCTTAAGCCACACATTAGCTGCACCTCGTGGCTGCAGCCAAA 2288
Db      2261 GCTCAGGTAAGAGAAGAGTCCCAAGATGCATTCAGGCTGGGCTGGAGCCAAAGCAGA 2320
Qy      2289 ACAAGAACTGGGTGT---TGAGTATTCATCACTAAGAACCC---AAATCCAGGGCACT 2342
Db      2321 GGAAGAACTAGCTACTGACCGAGCATTCAGAGCAAGGTCCTTTCAAAATCCAGGGCACA 2380
Qy      2343 CATATGTGAAGATAAAGAACCTCACTTCTTACTTCTTCCCAAAA--GAAGTGGGAAGA 2400
Db      2381 GCTGTGATGACAGAGAGCCCCACACCTCTTTAGTCTCTAAAGAGGAGATGGGGGTGA 2440
Qy      2401 ACCATCAAACTTTCTCCTGACTTACCAACACAGGAAACAGCAGGAGGGTGGCTCA 2460
Db      2441 ATCAATGGGATTTCTCCTGACTTAGCAACAGGAGGAGCAGGAGCAAGACAGCTCA 2500
Qy      2461 GGACTTA 2467
Db      2501 GGATTTA 2507

```

RESULT 9

RATRAB15X

LOCUS

DEFINITION

ACCESSION

VERSION

945 bp mRNA linear ROD 27-APR-1993
 RATRAB15X Sprague-Dawley (clone LRB9) RAB15 mRNA, complete cds.
 M83679.1 GI:206536

```

KEYWORDS  LMW GTP-binding protein.
SOURCE     Rattus norvegicus (strain Sprague-Dawley) (library: LAMBDA ZAPII)
           adult brain cDNA to mRNA.
ORGANISM   Rattus norvegicus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
           Rattus.
REFERENCE  1 (bases 1 to 945)
AUTHORS   Elferink,L.A., Anzai,K. and Scheller,R.H.
TITLE      rab15, a novel low molecular weight GTP-binding protein
           specifically expressed in rat brain
JOURNAL   J. Biol. Chem. 267 (9), 5768-5775 (1992)
MEDLINE   92210533
PUBMED    1313420
FEATURES   Location/Qualifiers
            1..945
             /organism="Rattus norvegicus"
             /strain="Sprague-Dawley"
             /db_xref="taxon:10116"
             /tissue_type="brain"
             /dev_stage="adult"
             /tissue_lib="LAMBDA ZAPII"
            1..945
             /gene="RAB15"
             /gene="RAB15"
             /codon_start=1
             /product="RAB15"
             /protein_id="AAA41995.1"
             /db_xref="GI:206537"
             /translation="MAKQYDVLFRLLIGDSVGKTCLLCRFTNPFHSHSTIGVD
             FMKTIIVDGIKVIQIWDYAGRYQITIKOYRRAQGIPLVDYDISSERYQHIMKW
             VSDYDEYAPGVQKILIGNKADEQKQVREGQQLAKYEMDFYETSACTNLNIKE
             SFTRLTELVLQAHKRLDGLTCAENLALALEBEDEKTEGPANSSKTCW"
BASE COUNT 222 a 259 c 292 g 172 t
ORIGIN
Query Match 17.7%; Score 577.6; DB 10; Length 945;
Best Local Similarity 88.7%; Pred. NO. 1e-137;
Matches 637; Conservative 0; Mismatches 79; Indels 2; Gaps 1;
Qy      21 TTCCCGCGCCCGCTGGCCCGCAGTCATCGCGAAGCAGTACGATGCTCTTCCGGCTGCTG 80
Db      196 TTCCCGCGCCCGCTGGCTGAGTCCCTGCCATCGCAACAGTACGATGCTATTTCCGGTACTG 255
Qy      81 CTGATCGGGGACTCCGGGTTGGCAAGACCTGCTGCTGCTGCGCGCTTCACGACACAGAG 140
Db      256 CTGATCGGGGACTCCGGGTTGGCAAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 315
Qy      141 TTCACCTCTCGCACATCTCCACCATCGGTGTGACTTTTAAGATGAAGACCATAGAGTA 200
Db      316 TTCCACTCTCGCATATCTCCACCATCGGTGTGACTTTTAAGATGAAGACCATAGAGTA 375
Qy      201 GACGSCATCAAGTCCGGATACAGATCTGGGACACTGCGAGGAGGAGGAGATACAGAGC 260
Db      376 GACGSCATCAAGTCCGGATACAGATTTGGGACACAGCGAGGAGGAGGATACAGACT 435
Qy      261 ATCAAAAGCAGTACTATCGCGGGGCCCGAGGGATATTTTGGTCTATGACATAGCAGC 320
Db      436 ATCAAAAGCAGTACTATCGCGGGGCCCGAGGGATATTTTGGTCTATGACATAGCAGC 495
Qy      321 GAGGCTCTTACCAGCACATCATGAAGTGGGTGAGTGCAGTGCAGTGCAGTGCAGTGCAG 380
Db      496 GAGGCTCTTATCAGCACATCATGAAGTGGGTGAGTGCAGTGCAGTGCAGTGCAGTGCAG 555
Qy      381 GCGTCCAGAGATCCTTTATTTGGGAATAAGCTGTATGAGGAGCAGCAACAGCAGTGGGA 440
Db      556 GCGTCCAGAGATCCTTCATCGGGGAATAAGCTGTACGAAGAACACAGAACGCGAGTGGG 615
Qy      441 AGAGAGCAAGGGCAGCAGCTGGCGAAGGAGTATGCGATGGACTTCTATGAACAAAGTGC 500
Db      616 AGAGAGCAAGGGCAGCAGCTGGCTAAGAGGATGTCGATGGAGTCTTCTACGAACAAGTGC 675

```


LOCUS	AC096084	162495 bp	DNA	linear	HTG 24-AUG-2002
DEFINITION	Rattus norvegicus clone CH230-16122, *** SEQUENCING IN PROGRESS				
ACCESSION	AC096084	***, 70 unordered pieces.			
VERSION	AC096084.6	GI:22164942			
KEYWORDS	HTG; HTGS; PHASE1.				
SOURCE	Norway rat.				
ORGANISM	Rattus norvegicus				
REFERENCE	1 (bases 1 to 162495)				
AUTHORS	Muzny,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhaq,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,J., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K., Harvey,X., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.B., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jollivet,A., Karpach,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindratne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mathew,S., McLeod,M., McNeill,T., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokemele,O., Okwuonu,G., Olarunpogun,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.				
TITLE	Direct Submission				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 162495)				
AUTHORS	Worley,K.C.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA				
REFERENCE	3 (bases 1 to 162495)				
AUTHORS	Rat Genome Sequencing Consortium.				
TITLE	Direct Submission				
JOURNAL	Submitted (24-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One				

Baylor Plaza, Houston, TX 77030, USA	On Aug 9, 2002 this sequence version replaced gi:21723229.
Center: Baylor College of Medicine	Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/	Contact: hgsc-help@bcm.tmc.edu
Center project name: GEJZ	Center clone name: CH230-16122
Center clone name: CH230-16122	Summary Statistics
Sequencing vector: Plasmid	Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap	Assembly version: 0.990329
Consensus quality: 116551 bases at least Q40	Consensus quality: 123470 bases at least Q30
Consensus quality: 123470 bases at least Q30	Consensus quality: 127758 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently consists of 70 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1247: contig of 1247 bp in length	1248
1347: gap of unknown length	1348
2634: contig of 1287 bp in length	2635
3987: contig of 1253 bp in length	3988
4087: gap of unknown length	4088
5288: contig of 1201 bp in length	5289
5388: gap of unknown length	5389
6866: gap of unknown length	6867
8143: contig of 1457 bp in length	8144
8243: gap of unknown length	8244
9249: contig of 1006 bp in length	9250
9349: gap of unknown length	9350
10754: contig of 1405 bp in length	10755
10854: gap of unknown length	10855
12161: contig of 1307 bp in length	12162
1261: gap of unknown length	1262
13639: contig of 1378 bp in length	13640
13739: gap of unknown length	13740
14946: contig of 1207 bp in length	14947
15046: gap of unknown length	15047
16342: contig of 1296 bp in length	16343
16442: gap of unknown length	16443
1772: contig of 1330 bp in length	17772
17872: gap of unknown length	17873
19007: contig of 1135 bp in length	19008
19107: gap of unknown length	19108
20796: contig of 1689 bp in length	20797
20896: gap of unknown length	20897
22145: contig of 1249 bp in length	22146
22445: gap of unknown length	22446
24330: contig of 2085 bp in length	24331
24430: gap of unknown length	24431
25589: contig of 1159 bp in length	25590
25689: gap of unknown length	25690
27349: contig of 1660 bp in length	27350
27449: gap of unknown length	27450
28577: contig of 1128 bp in length	28578
28677: gap of unknown length	28678
30291: contig of 1614 bp in length	30292
30391: gap of unknown length	30392
32126: contig of 1735 bp in length	32127
32226: gap of unknown length	32227
33343: contig of 1117 bp in length	33343

Baylor Plaza, Houston, TX 77030, USA	
On Aug 9, 2002 this sequence version replaced gi:21723229.	
----- Genome Center	
Center: Baylor College of Medicine	
Center code: BCM	
Web site: http://www.hgsc.bcm.tmc.edu/	
Contact: hgsc-help@bcm.tmc.edu	
----- Project Information	
Center project name: GEJZ	
Center clone name: CH230-16122	
----- Summary Statistics	
Sequencing vector: Plasmid;	
Chemistry: dye-terminator Big Dye; 100% of reads	
Assembly program: Phrap; version 0.990329	
Consensus quality: 116551 bases at least Q40	
Consensus quality: 123470 bases at least Q30	
Consensus quality: 127758 bases at least Q20	

* NOTE: Estimated insert size may differ from sequence length	
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).	
* NOTE: This is a 'working draft' sequence. It currently	
* consists of 70 contigs. The true order of the pieces	
* is not known and their order in this sequence record is	
* arbitrary. Gaps between the contigs are represented as	
* runs of N, but the exact sizes of the gaps are unknown.	
* This record will be updated with the finished sequence	
* as soon as it is available and the accession number will	
* be preserved.	
1	1247: contig of 1247 bp in length
* 1248	1347: gap of unknown length
* 1348	2634: contig of 1287 bp in length
* 2635	2734: gap of unknown length
* 2735	3987: contig of 1253 bp in length
* 3988	4087: gap of unknown length
* 4088	5288: contig of 1201 bp in length
* 5289	5388: gap of unknown length
* 5389	6586: contig of 1198 bp in length
* 6587	6686: gap of unknown length
* 6687	8143: contig of 1457 bp in length
* 8144	8243: gap of unknown length
* 8244	9249: contig of 1006 bp in length
* 9250	9349: gap of unknown length
* 9350	10754: contig of 1405 bp in length
* 10755	10854: gap of unknown length
* 10855	12161: contig of 1307 bp in length
* 12162	12261: gap of unknown length
* 12262	13639: contig of 1378 bp in length
* 13640	13739: gap of unknown length
* 13740	14946: contig of 1207 bp in length
* 14947	15046: gap of unknown length
* 15047	16342: contig of 1296 bp in length
* 16343	16442: gap of unknown length
* 16443	17772: contig of 1330 bp in length
* 17773	17872: gap of unknown length
* 17873	19007: contig of 1135 bp in length
* 19008	19107: gap of unknown length
* 19108	20796: contig of 1689 bp in length
* 20797	20896: gap of unknown length
* 20897	22145: contig of 1249 bp in length
* 22146	22245: gap of unknown length
* 22246	24330: contig of 2085 bp in length
* 24331	24430: gap of unknown length
* 24431	25589: contig of 1159 bp in length
* 25590	25689: gap of unknown length
* 25690	27349: contig of 1660 bp in length
* 27350	27450: gap of unknown length
* 27450	28577: contig of 1128 bp in length
* 28578	28677: gap of unknown length
* 28678	30291: contig of 1614 bp in length
* 30292	30391: gap of unknown length
* 30392	32126: contig of 1735 bp in length
* 32127	32226: gap of unknown length
* 32227	33343: contig of 1117 bp in length

* 33444	33443: gap of unknown length	
* 34474	34975: contig of 1532 bp in length	
* 34976	35075: gap of unknown length	
* 35076	37579: contig of 2504 bp in length	
* 37580	37679: gap of unknown length	
* 37680	39003: contig of 1324 bp in length	
* 39004	39103: gap of unknown length	
* 39104	40564: contig of 1461 bp in length	
* 40565	40664: gap of unknown length	
* 40665	41979: contig of 1315 bp in length	
* 41980	42079: gap of unknown length	
* 42080	43598: contig of 1519 bp in length	
* 43599	43698: gap of unknown length	
* 43699	45418: contig of 1720 bp in length	
* 45419	45518: gap of unknown length	
* 45519	47359: contig of 1841 bp in length	
* 47360	47459: gap of unknown length	
* 47460	48918: contig of 1459 bp in length	
* 48919	49018: gap of unknown length	
* 49019	50383: contig of 1365 bp in length	
* 50384	50483: gap of unknown length	
* 50484	52346: contig of 1863 bp in length	
* 52347	52446: gap of unknown length	
* 52447	52557: contig of 2811 bp in length	
* 52558	53557: gap of unknown length	
* 53558	57005: contig of 1648 bp in length	
* 57006	57105: gap of unknown length	
* 57106	59670: contig of 2565 bp in length	
* 59671	59770: gap of unknown length	
* 59771	61952: contig of 2182 bp in length	
* 61953	62052: gap of unknown length	
* 62053	64503: contig of 2451 bp in length	
* 64504	64603: gap of unknown length	
* 64604	65752: contig of 1149 bp in length	
* 65753	65853: gap of unknown length	
* 65853	68166: contig of 2314 bp in length	
* 68167	68466: gap of unknown length	
* 68267	70091: contig of 1825 bp in length	
* 70092	70191: gap of unknown length	
* 70192	72887: contig of 2696 bp in length	
* 72888	72987: gap of unknown length	
* 72988	74782: contig of 1795 bp in length	
* 74783	74882: gap of unknown length	
* 74883	77248: contig of 2366 bp in length	
* 77249	77348: gap of unknown length	
* 77349	79279: contig of 1931 bp in length	
* 79280	79379: gap of unknown length	
* 79380	83158: contig of 3779 bp in length	
* 83159	83258: gap of unknown length	
* 83259	85363: contig of 2105 bp in length	
* 85364	85463: gap of unknown length	
* 85464	88527: contig of 3084 bp in length	
* 88528	88627: gap of unknown length	
* 88628	90451: contig of 1824 bp in length	
Query Match		13.0%; Score 423.6; DB 2; Length 162495;
Best Local Similarity		61.8%; Pred. No. 1.5e-97;
Matches 1227; Conservative		0; Mismatches 664; Indels 95; Gaps 31;
QY 523	AGTCATTCACGGCTCTACAGAGCTGGTGTCTGCAGGCCCATAGGAGGAGCTGGAGGCC 582	
DB 125590	AGTCTTCACTCGTCTGACCGAGCTGGTGTCTGCAGGCTCACAGAAAGAGCTGGATGGTC 125531	
QY 583	TCGGGATGGTCCACCAATAGTTGGCACTGGCAGAGCTGGAGGAGGAGGGCAAC 642	
DB 125530	TCGGAACATGTGCCAGCAATAGTCTGCACCTGGCCGAGCTGGAGGAAGACGAACAA 125471	
QY 643	CCGAGGCCAGGAACTCTCCGAAACCTGTGTGTGTGA--GTCTGTGTGGGCACC 700	
DB 125470	CTGAAGCCAGCAAACTCTCAAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 125411	
QY 701	CCACAGCACACCCCTCTCCCTCAGGAGGCCGTGGGCAGACAGGAGGCCGGGCTTTG 760	
DB 124316	GGTGACAGCAGGGCGGTCCCTGTGAGGACATTTACCCAGACTCCTCCCTGCAGAGGTCA 124316	
DB 125410	CCCTCCCCCACTCCCTTCCTCAGGAGGCCCGTGGAGCA---GGAGTCTGGGCTTTG 125354	
QY 761	CCCTGCTGCTGCTCTCTCGTGTGATGACCTATTAGTATCAGTAGCCTACTCTCCCT 820	
DB 125353	TCCATC-ACGTCTCTTTTCAATTGAGATCTGTTGAAACCACTAGCTACT-CCCT 125296	
QY 821	GCCTGGCCCTGAGAGCGGCTGCTGTCTATCT-CAAGCAGCCCTGTGCCAGCCGCTCC 879	
DB 125295	GCCTGGCCCTGTGAGCAGCTTGTCTGTGTGCTCGGAGCAACCTGTGTGCCAGCCCTGCT 125236	
QY 880	ACCTGTGAGTGGTCTCTTTCAGCCTGTTTCCAGCCACACAGGCTGTACAGCCCAACG 939	
DB 125235	ATCCTGGAGTGGTAACTTTTCCACCTGTTTCCCGCCGAGGCTGTGTGACCC--- 125180	
QY 940	ATGTGCGCGAGCACTGTCTCACCATCCCGCACCCACAGACACACAGGCTGGAGT 999	
DB 125179	AGATGTGCTTCGACCATCTCACCAACCCAGGCTCTGGGACACAGCTGAGT 125120	
QY 1000	CCAGGCACCTTTCAGCTGCTCTTTCGCGTGCATGCTGTCTCTCTCTCTCTCTCTCT 1059	
DB 125119	CAAAAGCCACTTAAAGACTCTCTTTAGTGCATCTCGTGGGCTCTCTGCTTTCTCTCTT 125060	
QY 1060	CTTCCCTCCACTCTCTTCTCTGACCTCCCTCCGCTGGTTCGTATCAAAAGCTCT 1119	
DB 125059	CTTCTCTTCTCCAGTCCCTTCCCTTCCATGATTTTGGTCCAGCTCATCCAGCTCTCT 125000	
QY 1120	CAAAACCCGCTCCCGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1179	
DB 124999	C---CTCCCTAGCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 124943	
QY 1180	TATCAAGGGGATGAGCCAGGCTCGTGGGAGGTTCCACCTTGGATCCAGGAAGAAC 1239	
DB 124942	TGCCTAAGGAATGAGCCAGGCTC-----TGCCCCAGGAGGATG 124902	
QY 1240	CTCCACCTGCTGCTGGTGGTGGCCAAAGGTACAGGCTGTCTCTCTCTCTCTCTCTCT 1239	
DB 124901	GCTCACCTG-CTCTGGGCTGGGCTGAAGGCTA-AGGCTGTCTCTCTCTCTCTCTCTCT 124846	
QY 1300	CCCTGTGCTCTATGTGCCATGGCTGCTCTCCAGTGTGTGTGTGTGTGTGTGTGTGT 1359	
DB 124845	CTGCTGCTCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 124787	
QY 1360	GAGGTAGGAGGAAACAGCAACCGGGAGTCTCTGAGCTGGGCTGCTCTACCTTACC 1419	
DB 124786	GAGGTAGGAGGAAACACTATGAGGTG-GTCTCAAGCT-GGACTGCTCTCTCTCTCTCT 124729	
QY 1420	CATTCCCGCAGCAGCTTTGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1479	
DB 124728	TGCTCTTCTGCTGGAGTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 124669	
QY 1480	GCTCAGAGGAGGCTTTCAGAGAGGAAACAAATGAGGCTGCGCAGGATAAAGATC 1539	
DB 124668	GTCCAGAGGAGGACACAGCAGAGAAAGATGACAAAGATGGGTGGCGGGATTAAGTCA 124609	
QY 1540	ACCTCCATCTCTACTCCCATGCAAGTGAACAAATTTCTCTCACTGCTCTCTCTCTCT 1599	
DB 124608	--CTCTGTCTCTACTCTCCCATGCAAGTGTGACA--GCCTTTTCCACCGGTTCTG 124553	
QY 1600	TTTAAAGATGGACCAAGGCTGTGGTGTCTCCAGGCGCAGGAGAGCCCTGGGGTCTAG 1659	
DB 124552	ATTTAAAGATGGATCAAGGCTGTAGGACG-CAGAGGGCAAGGGGATCCCCAGGCTCA- 124495	
QY 1660	TGACACTGTCCAGGCCAACCATCTCTCCAAAGGGAGCATTTTGAATGAAGACTAG 1719	
DB 124494	-GACACTGTGTAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 124436	
QY 1720	CTCTATGTATCAGGTTAAGACCAAGG-AGAGCTGGCCAGGACAGCAGCTTTGCACAGC 1778	
DB 124435	CTCTTAGATGGCAGGTTAGGACAAAGATAGAGGGGGCTTTCAGAACAGCTAGTGTCT 124376	
QY 1779	AGAGGGGATGTAGCAACAGCAGGCGCTCTAGGCGCCCTCTTCCATTCTTAGTGAAGA 1838	
DB 124375	GGTGACAGCAGGGCGGTCCCTGTGAGGACATTTACCCAGACTCCTCCCTGCAGAGGTCA 124316	

Db 551 GAGAGCAGGGCAGCAGCTGGCTAAGGAGTACGGCATGGACTTCTACGAAACAGTGCCT 610

QY 502 GC 503

Db 611 GC 612

RESULT 14

AC124556/c

LOCUS

DEFINITION

AC124556

AC124556.2 GI:21539788

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AC124556 257226 bp DNA linear HTG 23-JUN-2002

Mus musculus chromosome UNK clone RP23-246K11, WORKING DRAFT

SEQUENCE, 15 unordered pieces.

AC124556

AC124556.2 GI:21539788

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 257226)

McPherson,J.D. and Waterston,R.H.

The sequence of Mus musculus clone

Unpublished

2 (bases 1 to 257226)

McPherson,J.D. and Waterston,R.H.

Direct Submission

Submitted (14-JUN-2002) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

3 (bases 1 to 257226)

McPherson,J.D. and Waterston,R.H.

Direct Submission

Submitted (23-JUN-2002) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

On Jun 23, 2002 this sequence version replaced gi:21426694.

----- Genome Center -----

Center: Washington University Genome Sequencing Center

Web site: http://genome.wustl.edu/gsc/index.shtml

Contact: submissions@watson.wustl.edu

----- Project Information -----

Center project name: M.BA0246K11

----- Summary Statistics -----

Sequencing vector: M13; 0%

Chemistry: Dye-primer ET; 0% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 249197 bases at least Q40

Consensus quality: 250905 bases at least Q30

Insert size: 214000; agarose-fp

Insert size: 255826; sum-of-contigs

Quality coverage: 9.36 in Q20 bases; agarose-fp

Quality coverage: 7.92 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently

consists of 15 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 5033: contig of 5033 bp in length

5034 5133: gap of unknown length

5134 14755: contig of 9622 bp in length

14756 14855: gap of unknown length

23382: contig of 8527 bp in length

23383 23489: gap of unknown length

23483 38499: contig of 15017 bp in length

38500 38599: gap of unknown length

46661: contig of 8062 bp in length

46761: gap of unknown length

46762 57176: contig of 10415 bp in length

57177 57276: gap of unknown length

57277 86426: contig of 29150 bp in length

86427 86526: gap of unknown length

86527 104351: contig of 17825 bp in length

104352 104451: gap of unknown length

104452 122682: contig of 18231 bp in length

122683 122783: gap of unknown length

122784 169668: contig of 46886 bp in length

169669 169769: gap of unknown length

169770 244701: contig of 74933 bp in length

244702 244801: gap of unknown length

244802 246147: contig of 1346 bp in length

246148 246247: gap of unknown length

246248 248024: contig of 1777 bp in length

248025 248125: gap of unknown length

248126 251795: contig of 3670 bp in length

251796 251895: gap of unknown length

251896 257226: contig of 5332 bp in length.

Location/Qualifiers

1. 257226

/organism="Mus musculus"

/db_xref="taxon:10090"

/chromosome="UNK"

/clone="RP23-246K11"

1. 5033

/note="assembly_name:Contig10"

5134 14755

/note="assembly_name:Contig11"

14856 23382

/note="assembly_name:Contig12"

23483 38499

/note="assembly_name:Contig13"

38600 46661

/note="assembly_name:Contig14"

46762 57176

/note="assembly_name:Contig15"

57277 86426

/note="assembly_name:Contig16"

86527 104351

/note="assembly_name:Contig17"

104452 122682

/note="assembly_name:Contig18"

122783 169668

/note="assembly_name:Contig19"

169769 244701

/note="assembly_name:Contig20"

244802 246147

/note="assembly_name:Contig21"

246248 248024

/note="assembly_name:Contig22"

248125 251794

/note="assembly_name:Contig23"

251895 257226

/note="assembly_name:Contig24"

BASE COUNT 68762 a 59441 c 58995 g 68614 t 1414 others

ORIGIN

Query Match 11.6%; Score 378.4; DB 2; Length 257226;

Best Local Similarity 61.2%; Pred. No. 7.8e-86;

Matches 1205; Conservative 0; Mismatches 596; Indels 169; Gaps 30;

QY 523 AGTCATTCACGCGTCTGACAGAGCTGTGCTGCAGGCCCATAGGAAGAGCTGGAAGGCC 582

Db 135389 AGTCCTTCACTTGGTCTGACGAGGCTGTGCTGCAGGCCCATAGGAAGAGCTGGAAGGCC 135330

QY 583 TCCGGATCGTCCAGCAATGAGTTGGCAGCTGGCAGAGCTGGAGAGGAGGCAAC 642

Db 135329 TCCGAACACGTCGCCAGCAACGAGCTCGCACCTGGCGGCGAGCTGGAGAGGAGGCAAC 135270

QY 643 CCGAGGGGCCGAGCAACTCTTCGAAACCTCTGTGTGCTGTGCTGTGTGGGGCACCCC 702

[illegible]

Db	134265	CGGGCTCCTAGAACCAAGGTTAGACGAAGAGCAGGAGGAGC-----TTCCGA	1342165
QY	1775	CAGCAGAGGGGAATGTAGCAACAGCAGGAGCGCTCTCTAGGCCCCCATCTCCATTTCTTAGGT	1834
Db	134217	GAACAGCGAGTGTGTGCGGACTACAGGSCATCAGAGTGACAAGTAATC-----	134168
QY	1835	AAGAAGAGCATTTCTCTAGACTCCAGCGGGAGGACTGAGCCTTAGCCTTACGCAACCAAG	1894
Db	134167	-----AACCACCTCTCTTTCTACTGAACCCAGCCTTAGGAGTAGG	134127
QY	1895	GTTCCTCTGGGACCCAAAGTTTATGGGAGAAGGGCAAGAGACTTCATGGGAAGAGAGAGG	1954
Db	134126	GTTTCCCCAGAGTCTGGGACACAGGCCCTTAGACAAGGAAGAGAGGTTGGGTA-----	134072
QY	1955	AAGGCCCTGGGTAGAAACGCTTGTGTCTCTTTTGGCCCTTTAAGACAAAGCCTCAT	2014
Db	134071	-----TGATGGAAATACTTGGTGTCTTCTGCTCTCC-----AGGACTCC	134032
QY	2015	CTTGCCCTCTACCTCTCATAGGCTTGAAGGTTTGCCTCAACACACTGTGGCTACAGGTGG	2074
Db	134031	CTCCCCCTGGTTCCTTATTTCTAAGTCGGCTTTGCCAACTGCTGTGTGGCTTACCAAGG	133972
QY	2075	AGGGAAGAGGACTCCTCTCCAGAGTGTATGTTTCAGGAAGTTCCTTTAAACCCCATATG	2134
Db	133971	AGGGAAGAGAGACACCTCTCCAGAGTCTATGCTCAGGAAGTTCCTTTAAACCC--TG	133914
QY	2135	GCCAAAGAGTAGCTGTGTAGGAGGCCCTTTAAAGACGGAACAGTAATTTA-----C	2185
Db	133913	GTTCAAAGAGTAGTCTGAAGGAGGCCCTGMAAACAAACAAACAAACAAACAAAGC	133854
QY	2186	CAGTCTACTGGGTTCTGTCCCAAGCTGCCAAGGTGGCGAGGCCCTAGGAAGAGGGTCA	2245
Db	133853	CAACCTCGGTAGGGTTCTATCTTACATTTCCCAAGGTGGGAGGAGCTCAGGTAAAGAAGA	133794
QY	2246	TTCTTAAGCCACATTAAGTGCACCTGCGTGGCTGCAGCCAAACAAAGAACTGGGTGT-	2304
Db	133793	GTCCCAAGATGCATCTCAAGCTGGGGCTGGAGCCAAAGCAGAGGAAGTACAGCACTG	133734
QY	2305	--TGAGTATTCATCACTAAGAAGC---AAATCAGGCACTCATATGTGAAGGATAG	2359
Db	133733	ACCAGGACATCAAGGACCAAGGCTCTTTCAATCAGGSCACAGCTGTGATCAGCAGAG	133674
QY	2360	AACCTCACTTCTTACTCTCTCCAAAA--GAAGTGGGGAAGAAACCATCAACCTTTCT	2417
Db	133673	CCCACACCTTTTAGTCTCTTAAAGGGAGATGGGGTGAATCAATGGGGATTTCTC	133614
QY	2418	CCTGACTTACCAAACAGGAAAAACAGCAGGAGGAGGCTGGCTCAGGACTTA	2467
Db	133613	CCTGACTTAGCAACAGGAGCAGGAGAGAGAGACAGCTCAGGATTTA	133564
RESULT 15			
AX341683			
LOCUS	AX341683	313 bp	DNA
DEFINITION	Sequence 1930 from Patent WO0196388.	linear	PAT 10-JAN-2002
ACCESSION	AX341683		
VERSION	AX341683.1	GI:18137665	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Jiang, Y., Harlocker, S.L. and Secrist, H.		
TITLE	Compositions and methods for the therapy and diagnosis of colon cancer		
JOURNAL	Patent: WO 0196388-A 1930 20-DEC-2001;		
	CORIXA CORPORATION (US)		
FEATURES	Location/Qualifiers		
	1..313		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	68 a	85 c	85 g 73 t 2 others

Result No.	Query #	Score	Match	Length	DB	ID	Description
C 1	2696.6	82.8	3240	22	AA882285		Human immune/haema
C 2	1568.2	48.1	1666	21	AA675813		Human ORF1368
3	1192.8	36.6	7924	24	ABK40069		Human chemically p
4	1192.8	36.6	7924	24	ABJ34138		Human immune syste
C 5	1160.8	35.6	7924	24	ABK40070		Human chemically p
C 6	1160.8	35.6	7924	24	ABJ34139		Human immune syste
7	1160.6	35.6	2021	22	AA527053		CDNA encoding nove
8	866.2	26.6	895	20	AA324400		Human bladder tumo
9	534.2	16.4	566	22	AA527472		CDNA encoding nove

PR 28-JUN-2000; 2000US-0214886.
 DR 30-JUN-2000; 2000US-021513E

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being predicted and is derived by analysis of the total score distribution.

Db 1180 GGAAGAACCAATCAAACTTCTCCTGACTTACCAAAACAGGAAACAGCAGGAGG 1121
QY GTGCTCAGGACTTAGGACAGGCTATAGCTTAGATGCTGGAAGCAAGAGCAGGACGAG 2512
Db 1120 GTGCTCAGGACTTAGGACAGGCTATAGCTTAGATGCTGGAAGCAAGAGCAGGACGAG 1061
QY AGTTGTAATCACTGGCTTAATGAGAAAGAGACAGCCTAACTCTAGGATGAAGCTGTGAC 2572
Db 1060 AGTTGTAATCACTGGCTTAATGAGAAAGAGACAGCCTAACTCTAGGATGAAGCTGTGAC 1001
QY TAGCTGGAGTGTCTTCTTGAAGATGGGACTCTTGGGTATCAAGACCTATGCCACATC 2632
Db 1000 TAGCTGGAGTGTCTTCTTGAAGATGGGACTCTTGGGTATCAAGACCTATGCCACATC 941
QY ACCTGGGGCTAGGAGTAGTGTATGACAGCCCTCAAGTCTGTCTTCAGCAGGACTT 2692
Db 940 ACCTGGGGCTAGGAGTAGTGTATGACAGCCCTCAAGTCTGTCTTCAGCAGGACTT 881
QY GAGAAGTTATATTGGGCAAGTGGCTCCAACTCTGTGACAGCATTTTCAGTTTCCCTGAAG 2752
Db 880 GAGAAGTTATATTGGGCAAGTGGCTCCAACTCTGTGACAGCATTTTCAGTTTCCCTGAAG 821
QY ATCAGGAGGCTGCATTCATCTTCTCTCTCTAGCCCTCAGGAAGAGACTAT 2812
Db 820 ATCAGGAGGCTGCATTCATCTTCTCTCTCTAGCCCTCAGGAAGAGACTAT 761
QY ATTTGACTGTACCTAGGGCTTCTGGAAGGAAACATGCAATCAGGATTTCTATAGACT 2872
Db 760 ATTTGACTGTACCTAGGGCTTCTGGAAGGAAACATGCAATCAGGATTTCTATAGACT 701
QY GATAGGCCCTATCCACAAGGGCCATGACTGGGAAAGGTATGGGAGCAGAGGAGATTTG 2932
Db 700 GATAGGCCCTATCCACAAGGGCCATGACTGGGAAAGGTATGGGAGCAGAGGAGATTTG 641
QY GATTTTATGGTGTACGTACCTCACTTAACTTTTGTGGCTGGGCACTGTCTTGAG 2992
Db 640 GATTTTATGGTGTACGTACCTCACTTAACTTTTGTGGCTGGGCACTGTCTTGAG 581
QY GCGCAGAGTTAAGCAGGCTGTCTGCTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 3052
Db 580 GCGCAGAGTTAAGCAGGCTGTCTGCTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 521
QY GTCCTGAGACTTCATCCAGCCCGCAGGACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3112
Db 520 GTCCTGAGACTTCATCCAGCCCGCAGGACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 461
QY TGTGACGGGTGAACCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3172
Db 460 TGTGACGGGTGAACCTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 401
QY CATCTATTTTAAACACAGATGTTTACAAATAAAGATTTTCAACACCACC 3223
Db 400 CATCTATTTTAAACACAGATGTTTACAAATAAAGATTTTCAACACCACC 350

RESULT 2

AACT75813
ID AACT75813 standard; cDNA; 1666 BP.

XX AC AACT75813;

XX DT 08-FEB-2001 (first entry)

XX DE Human ORFX ORF1368 polynucleotide sequence SEQ ID NO:2735.

XX KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
XX KW vulnerability; antiparkinsonian; nootropic; neuroprotective;
XX KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
XX KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
XX KW antiviral; antibacterial; antifungal; antineoplastic; antithyroid;
XX KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
XX KW neurodegenerative disorder; osteoarthritis; graft vs host disease;

KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.

XX Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

XX P-PSDB; AAB41604.

XX Novel nucleic acids and peptides derived from open reading frame X,

XX useful for treating e.g. cancers, proliferative disorders,

XX neurodegenerative disorders and cardiovascular disease -

XX Claim 5; Page 1978-1979; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnery;
XX antiparkinsonian; nootropic; neuroprotective;
XX osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
XX immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
XX antithyroid; and antianemic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating
XX pathological conditions associated with an ORFX-associated disorder. The
XX nucleic acids can be used to express ORFX proteins in gene therapy.
XX vectors. The proteins and nucleic acids may be used to treat cancers,
XX proliferative disorders, neurodegenerative disorders, osteoarthritis,
XX graft vs host disease, cardiovascular disease, diabetes mellitus,
XX hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
XX erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
XX bacterial or fungal infection, malaria, autoimmune disorders, asthma,
XX allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
XX nocturnal haemoglobinuria, antinflammatory disease; to enhance
XX coagulation; to inhibit thrombosis; and as a contraceptive.

XX SQ Sequence 1666 BP; 333 A; 529 C; 461 G; 341 T; 2 other;

Query Match 48.1%; Score 1568.2; DB 21; Length 1666;

Best Local Similarity 98.7%; Pred. No. 0;

Matches 1598; Conservative 0; Mismatches 3; Indels 18; Gaps 1;

QY 2 GCCCGCTGCCCGCCGCGAGTTCCCGGGCCCCCGCTGCCCGAGCAGTACGAGCAGTACGAG 61

DB 48 GCCCGCTGCCCGCCGCGAGTTCCCGGGCCCCCGCTGCCCGAGCAGTACGAGCAGTACGAG 107

QY 62 TGTGCTGTTCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 121

DB 108 TGTGCTGTTCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 167

QY 122 CCGCTTCCACGACAGCAGTTCCTCTCTCCACATCTCCACCATCTGCTGCTGCTGCTGCTGCTG 181

DB 168 CCGCTTCCACGACAGCAGTTCCTCTCTCCACATCTCCACCATCTGCTGCTGCTGCTGCTGCTG 227

each case at least one base sequence having a length of 9 nucleotides which hybridizes to or is identical to a chemically pretreated DNA of genes associated with pharmacogenomics and their complements, arranged in an array for analysing diseases associated with the methylation state (CpG) and/or detecting SNPs (single nucleotide polymorphisms) of the 87 sequences. The oligomers may also be used as PCR primers. The set of 87 nucleic acids and their complements is useful for diagnosis and therapy of solid tumours and cancer. The present sequence represents one the 87 DNA sequences or its complement.
Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 7924 BP; 2087 A; 83 C; 2037 G; 3717 T; 0 other;

Query Match		36.6%;	Score 1192.8;	DB 24;	Length 7924;
Best Local Similarity		75.8%;	Pred. No. 5e-279;		
Matches 1476;		Conservative 0;	Mismatches 472;	Indels 0;	Gaps 0;
QY	1276	GGTGCTTCCTCTCTCCACCACCCACCTGCTCCATGTCCTCCATGGGCGCTGCTCCCTCC	1335		
DB	1	GGTGCTTCCTCTCTCTCCACCACCCACCTGCTCCATGTCCTCCATGGGCGCTGCTCCCT	60		
QY	1336	AGTGACCTGCGAAAGTCGAGCATCGAGGTAGAGGGAACACAGCACCGGGAGTCTCGA	1395		
DB	61	AGTGATTTTCGAAAGTCGAGCATCGAGGTAGAGGGAACACAGCACCGGGAGTCTCGA	120		
QY	1396	GCTCGGGGTCGCTACCTCTACCCATCCCGACAGAGCTTCGCCCTGCTGGCTGC	1455		
DB	121	GGTGGGCTGCTTCCTCTCTCTCCACCACCCACCTGCTCCATGTCCTCCATGGGCGCT	180		
QY	1456	CGCCCTGCTCTCTCTCCACCACCCACCTGCTCCATGTCCTCCATGGGCGCTGCTCCCT	1515		
DB	181	TCGCTTCTCTCTCTCTCCACCACCCACCTGCTCCATGTCCTCCATGGGCGCTGCTCCCT	240		
QY	1516	GAGGGGTGCGAGGATTAAGATCAGCTCCATCTCTACCTCCATGTCAGGAGGAAACAAAT	1575		
DB	241	GAGGGGTGCGAGGATTAAGATCAGCTCCATCTCTACCTCCATGTCAGGAGGAAACAAAT	300		
QY	1576	ATTTCCTCTCCACTGCTCCACAAATTAAGATGTCGACCAAGGCTGTGGTACTCCAG	1635		
DB	301	ATTTCCTCTCCACTGCTCCACAAATTAAGATGTCGACCAAGGCTGTGGTACTCCAG	360		
QY	1636	GGCAAGGAGAGCCCTGGGCTCAGTCAGCTGTCAGGCCAACCATGTCACCTCCACAAAGG	1695		
DB	361	GGGTAAGGAGAGCTTTGGGCTAGTCATGTCAGGAGGAAATGTCAGGAGGAAATGTCAG	420		
QY	1696	GAGCATTTGGAATGAAGGACTAGCTCTATGTCAGGTTAAGAGCAAGGAGAGCTGG	1755		
DB	421	GAGTATTTGGAATGAAGGACTAGCTCTATGTCAGGTTAAGAGCAAGGAGAGTGG	480		
QY	1756	CCAGGACAGCAGTTTTCACAGAGGGAATGTAGCAACAGAGGCGCTCCTAGGCC	1815		
DB	481	TTAGGATAGTAGTTTGTATAGTAGAGGGAATGTAGTAAATAGTATTTTATAGGTTT	540		
QY	1816	CATCTTCCATTTCTAGTTAAGAGAGCTTTCTCAGACTCCCGCGGAGGACTGACC	1875		
DB	541	TATTTTCTTCTAGTTAAGAGAGCTTTCTCAGACTCCCGCGGAGGACTGAGT	600		
QY	1876	CTAGCTTCAGAACCAAGCTTCTCTGGGACCCAAAGTTTATGGGAGAGGCAAGAC	1935		
DB	601	TTAGTTTCTAGTAATTAAGGTTTCTCTGGGACCCAAAGTTTATGGGAGAGGCAAGAT	660		
QY	1936	TTTATGGGAGAGAGGAGGAGGCGCTGGGTAGAAACGGTGGTGTCTCTCTTTGGCC	1995		
DB	661	TTTATGGGAGAGAGGAGGAGGCGCTGGGTAGAAACGGTGGTGTCTCTCTTTGGTT	720		
QY	1996	TTTATAGCAAGAGGCTCACTTCCCTCTACCTCTCTAGAGCTTCTAGGGTTTGGCAACC	2055		
DB	721	TTTATAGCAAGAGGCTTATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	780		
QY	2056	ACACTGTGGCTACAGGTGGAGGGAAGGAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2115		

DB	781	ATATGTGGTATATAGTGGAGGGAAGAGGATTTTTTTTTTAGAGTGTATGTGTAGGAA	840		
QY	2116	GTTCCTTTTAAACCCCATATGCCCAGAGTAGCTCGTAGGAGGCCCTTTAAAGACGGAA	2175		
DB	841	GTTCCTTTTAAATTTATATAGTTTAAAGTAGTGTCTGAGGAGGTTTTTAAAGACGGNATA	900		
QY	2176	AGTAATTTACCACTTCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2235		
DB	901	AGTAATTTATTTAGTTTATTTAGTTTATTTAGTTTATTTAGTTTATTTAGTTTATTTAG	960		
QY	2236	AAGAGGTCTATCTTAAGCCACATAGCTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2295		
DB	961	AAGAGGTCTATCTTAAGCCACATAGCTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1020		
QY	2296	ACTGGTGTCTGAGTATTCATCAAGAACCAAAATCCAGGCGACATCATATGTGAAGGA	2355		
DB	1021	ATTCGGTGTCTGAGTATTCATCAAGAACCAAAATCCAGGCGACATCATATGTGAAGGA	1080		
QY	2356	TAAGAACCTCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2415		
DB	1081	TAAGAACCTCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1140		
QY	2416	CTCCTGACTTACCAAAACAGGAAACAGCAGGAGGCTGCTCAGGACTTTAGGACAGG	2475		
DB	1141	TTTTTGTATTTAATTTAGGAAATAGTAGAGAGGCTGCTTAGGATTTAGGAGTAGG	1200		
QY	2476	GTATAGCTTAGATGGTGGAAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	2535		
DB	1201	GTATAGCTTAGATGGTGGAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1260		
QY	2536	GAAAGGAGAGCAGCTACTCTAGGATGAGCTGTCAGTGGCTGAGTGGCTGCTCTCTCTGAA	2595		
DB	1261	GAAAGGAGAGCAGCTACTCTAGGATGAGCTGTCAGTGGCTGAGTGGCTGCTCTCTCTGAA	1320		
QY	2596	GATGGGACTCTCTGGGTATCAAGACCTATGCCACATCACACTGGGGCTAGGGAAGTAGGT	2655		
DB	1321	GATGGGACTCTCTGGGTATCAAGACCTATGCCACATCACACTGGGGCTAGGGAAGTAGGT	1380		
QY	2656	GATGGGAGCCCTCAAGCTGCTCTCAGCCAGGAGCTTCAGAAAGTTATATTTGGGAGGAGG	2715		
DB	1381	GATGGTACTTTTAAAGTTGTTTGTAGTTAGGATTTTGAAGTTATATTTGGGTAGTGGT	1440		
QY	2716	TCCATCTGTCGACCAAGTATTTAGCTTTTCCCTGAAGATCAGGAGGCTGCTCATTTG	2775		
DB	1441	TTTATTTGTCGATGATGATTTTGTAGTTTGTGAGATTTAGTGGGTGTTTATTTATG	1500		
QY	2776	TCTTCTCTCTCTAGCCCTCAGGAAAGAGGAGCTATATTTGTACTGTACCTAGGGGT	2835		
DB	1501	TTTTTTTTTTTTTTAGTTTATTTAGGAAAGAGGATTTATTTGTATTTTATTTAGGGGT	1560		
QY	2836	CTGGAAGGAGAAACATGGAATCAGGATTTCTATAGACTCATAGGCTTATCCACAGGCC	2895		
DB	1561	TTGGAAGGAGAAATGGAATGGAATTTATAGATTTATAGTGTATTTATTTATAGGGTT	1620		
QY	2896	ATGACTGGGAAAGGATATGGGAGCAAGGAGATTTGGATTTTATAGGTGACGTACGCT	2955		
DB	1621	ATGATTTGGAAAGGATATGGGAGTAGAGGAGATTTGGATTTTATAGGTGAGTTACGTT	1680		
QY	2956	CACCTAAACCTTTTGGTGGCCCTGGGCAATGCTTGGAGGCCAGAGCTGTATAGGAGCTCT	3015		
DB	1681	TATTTTAAATTTTGGTGGCTATGTTTGGAGGTTTATTTGTATTTTATTTATTTAGGTTT	1740		
QY	3016	GCTGGCCTTTTCTCTGCTCACCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3075		
DB	1741	GTGGTGTGTTTATTTCTGTTATTTTGTATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG	1800		
QY	3076	AGGCGGCGCCTGCTCTCTGAGCCCTCCATCTATCTCTCTGCTGCTGCTGCTGCTGCTGCTG	3135		
DB	1801	AGGTCAGTTATTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT	1860		
QY	3136	TGTGCTCGGGTCCATATATGAAATGTGTAGCAGGCTTCATCTATTTTAAACACAGATGTT	3195		

```
Db 1861 TGTGTTTCGGGTTATATATGAATGTTGAGTAGCGGTTATTTATTTAAATATAGATGTT 1920
QY 3196 TACAAAATAAGATTATTTCAACACC 3223
Db 1921 TATAAAATAAGATTATTTAAATATC 1948

RESULT 4
ID ABL34138
AC ABL34138 standard; DNA: 7924 BP.
XX
XX ABL34138;
XX
XX 26-MAR-2002 (first entry)
XX
XX Human immune system associated gene SEQ ID NO: 2111.
XX
XX Human; Immune system disease; cytosine methylation; antiasthmatic;
XX antiarteriosclerotic; antiandemic; cytosine; nontropic;
XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
XX antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
XX antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
XX acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
XX neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
XX gene; ds.
XX
XX Homo sapiens.
XX
XX WO200200928-A2.
XX
XX 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP07537.
XX
XX 30-JUN-2000; 2000DE-1032529.
XX
XX 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIC-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful
XX for diagnosis and treatment of diseases associated with abnormal
XX cytosine methylation -
XX
XX Claim 1; SEQ ID NO 2111; 32pp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
XX genes which are modified by the methylation of cytosines. The sequences
XX can be used in the diagnosis and treatment of immune system disorders,
XX including eye diseases such as retinopathy, neovascular glaucoma and
XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX diseases. The present sequence is a gene of the invention.
XX
XX Sequence 7924 BP; 2087 A; 83 C; 2037 G; 3717 T; 0 other;
XX
XX Query Match 36.68; Score 1192.8; DB 24; Length 7924;
XX Best Local Similarity 75.8%; Pred. No. 5e-279;
XX Matches 1476; Conservative 0; Mismatches 472; Indels 0; Gaps 0;
XX
QY 1276 GGTGCTTCTCCCTTCCCCACCCACCTGCTCCCTCATGTGCGATGGCGCTGCTCC 1335
Db 1 GGTGTTTTTTTTTTTTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTT 60
QY 1336 AGTGACCTGGAAAGTGGAGCATCGAGGTAGGAGGAAACAGCAACCGGGAGTCTCGA 1395
Db 61 AGTGATTGCGAAAGTGGAGTATCGAGGTAGGAGGAAACAGCAATTTAGGAGTTTCGA 120
QY 1396 GCCTGGGGCTGCCTACCTCTACCCATTCACCCAGACAGCTTTGCCCTTGTGCTGTC 1455
```


Db 7562 AAAAAAAAAAACCCTAAATCAATACACATATCAACCAACCACTACACTCCACAAAAA 7503
QY 1698 GCATTGGAAATGAAGACTAGCTCCTATGATATCAGTTTAAGAGCAAGGAGAGCTGGCC 1757
Db 7502 ACATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7443
QY 1758 AGGACAGAGTTTGCACAGCAGAGGGATGTAGCAACAGAGAGGCTCCTAGGCCCA 1817
Db 7442 AAAAAACAATTTACACACAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 7383
QY 1818 TCTTCCATTTCTTAGTGAAGAGCATTTCTCAGACTCCAGGCGGAGGACTGAGCCT 1877
Db 7382 TCTTCCATTTCTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7323
QY 1878 AGCCTTCAGCAACAAAGTTCTCCTGGGACCCAAAGTTTATGGGAGAGGCGCAAGACTT 1937
Db 7322 AACCTTCAACAACCAAAATTTCTCTAAACCCCAAAATTTATAAAAAAATAAATAA 7263
QY 1938 CATGGGAGAGAGAGGAGGCTGGGTAGAACGCTTGGTCTGTCTCTTTGGCCTT 1997
Db 7262 CATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7203
QY 1998 TAAGACAAAGCGCTCATTTGGCCTCTACCTCTGATAGGCTTGAAGGTTTGCACACAC 2057
Db 7202 TAAACAAACGCTCATTTACCTCTACCTCTCTATTAACCTTAAATAATTACACACAC 7143
QY 2058 ACTGTGGCTACAGGTGAGGAGAGAGGACTCTCTCTCCAGAGTGCTATGTTCCAGGA 2117
Db 7142 ACTATAACTACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7083
QY 2118 TTCTTTAACCCCATATGGCCCAAGAGTAGCTCTGTAGGAGCCCTTTAAAGACGAGACAG 2177
Db 7082 TTCTTTAACCCCATATTAACCCCAAAATAAATCTGTAAATAAATAAATAAATAAATAA 7023
QY 2178 TAATTTACCAATTTCTACTGGGTTCTGTCACACCGTCCCAAGGTGGCGAGGCTAGGAA 2237
Db 7022 TAATTTACCAATTTCTACTAAATTTCTACCCCGTCCCAATTAACGAACTTAAATA 6963
QY 2238 GAGGCTCATTTCTTAAGCCACACATATAGTGCATGCGGTGGCTGCAGCCCAAAACAA 2297
Db 6962 AAAAATCATTTCTTAAACCCACATTAATCACTAGCTAGCTAACTACAAACCAAAAC 6903
QY 2298 TGGGTGTGAGTATTCATCACTAAGAACCAAAATCCAGGCACTCATATGTGAAGATA 2357
Db 6902 TAAATATTAATATTCATCACTAATAAACCATAAACCATAAACCATAAATAAATAA 6843
QY 2358 AGAAGCTCATTTCTTACTCTCCAAAAAGAGTGGGAAAGAACATCAAACTTTCTCT 2417
Db 6842 AAAAGCTCATTTCTTACTCTCCAAAAAATAAATAAATAAATAAATAAATAAATAA 6783
QY 2418 CTTGACTTACCAACCAAGGAAACAGCAGGAGGGTGGCTCAGGACTTAGGACAGGGT 2477
Db 6782 CTTAACTTACCAACCAAAACCAACAAAAAATAAATAAATAAATAAATAAATAA 6723
QY 2478 ATAGCTTAGATGTGGAACCAAGAGAGCAGAGAGTTGTAATCACTGGCTTAATGAGA 2537
Db 6722 ATAACCTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6663
QY 2538 AAAAGGACAGCTAACTTCTAGGATGAAGCTGTGAGCTAGGCTTGGCTTCTTGAAGA 2597
Db 6662 AAAAAACAACCTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6603
QY 2598 TGGGACTCTTGGGTATCAAGACCTATGCCACATCACTAGGCGCTAGGGAAGTAGTGA 2657
Db 6602 TAAAGCTCTTAAATATCAAAAGCTATACCACATCACTTAAAGCTTAAATAAATAA 6543
QY 2658 TGGCAGCCCTCAAGTCTGTCTTACGCCAGGAGCTTGNAGAGTTATATTGGGAGTGCTC 2717
Db 6542 TACCAACCCCTCAATCTTCTTCAACCAAAACCTTAAAAAATTTATATTAACCAATRACT 6483
QY 2718 CAATCTGTGACAGATTTTCACTTTCCCTGAAGATCAGGCGGGTGCCATTCATTGTC 2777
Db 6482 CAATCTAATAACCAATATTTCACTTTCCCTTAAATCAACAAATAACCAATTCATTATC 6423

QY 2778 TTTCTCTCTAGCCCTCAGAAAGAGGACTATATTTGTACTGTACCTAGGGTTCT 2837
Db 6422 TTTCTCTCTTACCCCTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 6363
QY 2838 GGAAGGAAAAACATGGATCAGGATCTATAGACTGATAGCCCTATCCACAGGGCCAT 2897
Db 6362 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6303
QY 2898 GACTGGAAAAAGGTATGGGAGCAGAGGAGAAATTTAGGGTGCGAGCTACGCTCA 2957
Db 6302 AACTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6243
QY 2958 CCTAAACTTTTGGTGGCCTGGGCGCATGTCTTGAGGCCAGACTGTTTAAAGCAGCTCTGC 3017
Db 6242 CCTAAACTTTTAAATAACCTTAAACATATCTTAAACCAAACTATTTAAACCAAACTCTAC 6183
QY 3018 TGCCCTGTTTACTCGTCGTCACCACTCTGCACCTGCTGCTGTCTGTGAGACTCCATCCAGCCCGAG 3077
Db 6182 TAACCTTATTTACTCGTCGTCACCACTCTACACTATCTTAAACCTTAAACCTTAAACCTTAAAC 6123
QY 3078 GCAGCCGACCTGCTGAGCGCTCCACTATCTCCTGTGAGCGGTGAACCTTCGTGTACTG 3137
Db 6122 ACAGCCGACCTACTCTTAAACCTTAACTTCTCCTATAGGATTAACCTTCGTATACTA 6063
QY 3138 TGCTCGGGTCCATATATGAATTTGTGAGCAGGGTTCATCTATTTTAAACACAGATGTTA 3197
Db 6062 TATCTCGAATCCATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6003
QY 3198 CAAATTAAGATTTATTTCAACCAACCAAAAAAATAAATAAATAAATAAATAAATAA 3257
Db 6002 CAAATTAAGATTTATTTCAACCAACCAAGATATACTACCTAATAAATAAATAAATAA 5943

RESULT 6
ABL34139/C
ID ABL34139 standard; DNA; 7924 BP.
XX ABL34139;
AC AC
DT 26-MAR-2002 (first entry)
XX Human immune system associated gene SEQ ID NO: 2112.
DE Human; immune system disease; cytosine methylation; antiasthmatic;
XX antiarteriosclerotic; antianemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX Homo sapiens.
OS
XX WO20020928-A2.
PN
XX 03-JAN-2002.
PD
XX 02-JUL-2001; 2001WO-EP07537.
PF
XX 30-JUN-2000; 2000DE-1032529.
PR
XX 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K;
PI WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -

XX	Claim 1; SEQ ID NO 2112; 32pp + Sequence Listing; German.	XX
XX	CC The present invention provides a number of human immune system associated	XX
XX	CC genes which are modified by the methylation of cytosines. The sequences	XX
CC	CC can be used in the diagnosis and treatment of immune system disorders,	CC
CC	CC including eye diseases such as retinopathy, neovascular glaucoma and	CC
CC	CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid	CC
CC	CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,	CC
CC	CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel	CC
CC	CC diseases. The present sequence is a gene of the invention.	CC
XX		XX
XX	Query Match 35.6%; Score 1160.8; DB 24; Length 7924;	XX
XX	Best Local Similarity 74.1%; Pred. No. 3e-271;	XX
XX	Matches 1468; Conservative 0; Mismatches 512; Indels 0; Gaps 0;	XX
QY	1278 TGCTTCTCTCTCTCCGCCACCCGACGCTGCTCCCTCATGTGCGATGGGCTCCCTCCCGAG 1337	QY
DB	7922 TACTTCTCTCTCTCCGCCACCCGACGCTGCTCCCTCATGTGCGATGGGCTCCCTCCCGAA 7863	DB
QY	1338 TGACCTGCGAAGTGGAGCATCGAGGTAGGAGGGAACAGCAACCGGGGAGTCTCTCGAGC 1397	QY
DB	7862 TAACCTACGAAATATAAATCGAATATAAATAAAGACGACAAATAATCTCTCGAAC 7803	DB
QY	1398 CTGGGCTGCGCTTACCTCTACCCATTCGCCGACGAGCTTGGCCCTTGGCTGGCTGCC 1457	QY
DB	7802 CTAAACTACCTTACCTCTACCCATTCGCCGACGAGCTTGGCCCTTGGCTGGCTGCC 7743	DB
QY	1458 GCCTGCTCTTTGGGGAAGTCTGAGCTCAGAGGAGGTGCTTCAGAGGAAGAAACAAATGA 1517	QY
DB	7742 GCCTACTCTTTAAAAAATCTAACTCAAAACAAATCTCTCAAAATAAATAAATAAATAA 7683	DB
QY	1518 GGGGTGCGAGGATAAAGTCTGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1577	QY
DB	7682 AAAATACAAATAAATAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7623	DB
QY	1578 TTCTCTCCACTGCTCCCAATTTAAAGATGTGGACCAAGGCTCTGGGTACTCCAGGG 1637	QY
DB	7622 TTCTCTCCACTGCTCCCAATTTAAATAATAAATAAATAAATAAATAAATAAATAAATAA 7563	DB
QY	1638 GCAAGGAGGCTGGGTGAGTGTGACATCTCAGCGCAACCATGACATCCCAAGGGGA 1697	QY
DB	7562 ACAAAAAAACCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7503	DB
QY	1698 GCATTTGGAATGAAGGACTAGTCTCTATGTATCAGGTTTAAGCAAGGAGAGAGTGGCC 1757	QY
DB	7502 ACATTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7443	DB
QY	1758 AGGACAGAGAGTTTGACAGCAGAGGGGAATGTAGCAACAGCAGGCGCTCTAGGCCCA 1817	QY
DB	7442 AAAAACACAATTTACACACAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7383	DB
QY	1818 TCTTCCATTTCTTAGGTAAAGAGCATTTCTCTCAGACTCCCGCGGAGGAGTCAAGCT 1877	QY
DB	7382 TCTTCCATTTCTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7323	DB
QY	1878 AGCCTTCAGCAACCAAGTTCTCTCTGGACCCCAAGTTTATGGGAGAGGGCAAGACTT 1937	QY
DB	7322 AACCTTCAACCAACCAATTTCTCTTAAACCCCAATTTTAAATAAATAAATAAATAAATA 7263	DB
QY	1938 CATGGGAAGAGAGAGGAGGCGCTGGGTAGAAAGCTTGGTGTCTCTCTCTCTCTCTCT 1997	QY
DB	7262 CATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 7203	DB
QY	1998 TAAGACAAGCGCTCATCTTGGCTCTTACCTCTCTGATAGGCTTGGAGGTTTGGCAACAC 2057	QY
DB	7202 TAAACAAAGCGCTCATCTTACCTCTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7143	DB
QY	2058 ACTGTGCTACAGTGGAGGGAAGAGACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2117	QY
DB	7142 ACTATAACTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7083	DB
QY	2118 TTCTTTAAACCCCATATGCCCAAGAGTAGCTCGTAGGAGGCGCTTTTAAAGACGGAACAAG 2177	QY
DB	7082 TTCTTTAAACCCCATATGCCCAAGAGTAGCTCGTAGGAGGCGCTTTTAAAGACGGAACAAG 7023	DB
QY	2178 TAATTTACCAAGTTCTACTGGGTTCTCTGCCACCGCTCCCAAGGTGGGCGAGCTAGGAA 2237	QY
DB	7022 TAATTTACCAATTTCTACTAAAAATTTCTCCACCGCTCCCAATAAAGCAACCTAAAAA 6963	DB
QY	2238 GAGGTCATTTTAAAGCCACACATTTAGCTGCACTGCGTGGCTGAGCCAAAAACAAGAAC 2297	QY
DB	6962 AAAAATCATTTTAAACCCACACATTTAATACATACGTACGTAACTACACCAACCAAAAAA 6903	DB
QY	2298 TGGGTGTTGAGTATTTCATCACTAAGAACCAAAATCCAGGCGACTCATATGTGAAGATA 2357	QY
DB	6902 TAAATATTAATATTTATCATCACTAATAAACCAAAATCCAAACACTCATATATAAATAA 6843	DB
QY	2358 AGAACCTCACTCTCTTACTCTCCAAAAAGAGTGGGGAAGAACCACTCAAAACCTTTCTCT 2417	QY
DB	6842 AAAACCTCACTCTCTTACTCTCCAAAAATAAATAAATAAATAAATAAATAAATAAATAA 6783	DB
QY	2418 CTTGACTTACCAAAACCCAGGAAACAGCAGGAGGTGGCTCAGGACTTAGGGACAGGT 2477	QY
DB	6782 CTTAATTTACCAACCCAAAAAACCAACAAAAATAAATAAATAAATAAATAAATAAATAA 6723	DB
QY	2478 ATAGCTTAGATGGTGGAAAGCAAGAGAGCAGGAGGTGTGAATCTAGTCTATGAGA 2537	QY
DB	6722 ATAACTTAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6663	DB
QY	2538 AAGGAGACAGCTTAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2597	QY
DB	6662 AAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6603	DB
QY	2598 TGGGACTCTCTGGGTATCAAGACCTATGCCACATCACACTGGGGCTAGGGAAGTAGGTGA 2657	QY
DB	6602 TAAACTCTTAAATATCAAAACCTATACCACATCACACTAAATAAATAAATAAATAAATA 6543	DB
QY	2658 TGGCAGCCTCAAGTCTGCTTTCAGCCAGGAGTGTGAGAGTGTATATTGGGCGAGTGGCT 2717	QY
DB	6542 TACCAACCTCTCAATCTATCTTCAACCAAAAAATTTAAATAAATAAATAAATAAATAA 6483	DB
QY	2718 CAATCTGTGGACAGTATTTTCAAGTCTTCCCTGAAAGTATCAGGAGGCTGCATCTATTCTC 2777	QY
DB	6482 CAATCTATAACCAATATTTCAACTTTCCCTTAAATCAAAATAAATAAATAAATAAATAA 6423	DB
QY	2778 TTTCTCTCTAGCCCTCTCAGGAAAGAGACTATATTGTGTACTGTACCTTAGGGGTTCT 2837	QY
DB	6422 TTTCTCTCTTAAACCCCTCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6363	DB
QY	2838 GGAAGGAAAAACATGGAATCAGGATCTATAGACTGTAGGCGCTTATCCAAAGGCCAT 2899	QY
DB	6362 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6303	DB
QY	2898 GACTGGGAAAAAGTATGGGAGCAGAGGAGAAATTTGGGATTTTAGGTGCAGCTACGCTCA 2957	QY
DB	6302 AACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 624	DB
QY	2958 CCCTAAACTTTTGGTGGCCTGGGGCATGCTTGAG	

PD 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US01312.
PF 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226581.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234277.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0239960.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465460/50.
XX DR

QY 2822 GTACCTAGGGTTCTGGAAGGAAAAACATGGAATCAGGATTTCTATAGACTGATAGGCC 2881
 |||||
 Db 241 GTACCTAGGGTTCTGGAAGGAAAAACATGGAATCAGGATTTCTATAGACTGATAGGCC 300
 |||||
 QY 2882 TATCCACAGGCCATGACTGGGAAAGGATATGGAGCAGAGGAGAAATTTGGATTTAG 2941
 |||||
 Db 301 TATCCACAGGCCATGACTGGGAAAGGATATGGAGCAGAGGAGAAATTTGGATTTAG 360
 |||||
 QY 2942 GGTGAGCTACCTCACCTAACTTTTGGTGGCCTGGGCAATGCTTTGAGGCCCAGACT 3001
 |||||
 Db 361 GGTGAGCTACCTCACCTAACTTTTGGTGGCCTGGGCAATGCTTTGAGGCCCANACT 420
 |||||
 QY 3002 GTTAAGCAGGCTCTGCTGGCCCTTTTACTGCTACCCACTCTGACCTGCTGCTTTGAGA 3061
 |||||
 Db 421 GTTAACAGGCTCTGCTGGCCCTTTTACTGCTACCCACTCTGACCTGCTGCTTTGAGA 480
 |||||
 QY 3062 C 3062
 |||||
 Db 481 C 481

RESULT 11

ABL6692/C

ID ABL66992 standard; DNA; 463 BP.

XX ABL66992;

XX

XX 15-MAY-2002 (first entry)

XX Thyroid cancer related gene sequence SEQ ID NO:5329.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.

XX Homo sapiens.

XX

XX WO200194629-A2.

XX

XX 13-DEC-2001.

XX

XX 30-MAY-2001; 2001WO-US10838.

XX

XX 05-JUN-2000; 2000US-209473P.

XX

XX 05-JUN-2000; 2000US-209531P.

XX

XX 18-SEP-2000; 2000US-233133P.

XX

XX 18-SEP-2000; 2000US-233617P.

XX

XX 20-SEP-2000; 2000US-234009P.

XX

XX 20-SEP-2000; 2000US-234034P.

XX

XX 20-SEP-2000; 2000US-234052P.

XX

XX 22-SEP-2000; 2000US-234509P.

XX

XX 22-SEP-2000; 2000US-234567P.

XX

XX 25-SEP-2000; 2000US-234923P.

XX

XX 25-SEP-2000; 2000US-234924P.

XX

XX 25-SEP-2000; 2000US-235077P.

XX

XX 25-SEP-2000; 2000US-235082P.

XX

XX 25-SEP-2000; 2000US-235134P.

XX

XX 25-SEP-2000; 2000US-235280P.

XX

XX 26-SEP-2000; 2000US-235637P.

XX

XX 26-SEP-2000; 2000US-235638P.

PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 XX (AVAL-) AVALON PHARM.

PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

PI Soppet DR, Weaver Z;

XX

XX WPI; 2002-188264/24.

XX

PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set

PS Claim 1; SEQ ID 5329; 44pp; English.

XX

XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilm's tumour.

SQ Sequence 463 BP; 116 A; 108 C; 110 G; 129 T; 0 other;

Query Match 13.5%; Score 438.8; DB 24; Length 463;

Best Local Similarity 99.3%; Pred. No. 1.7e-96;

Matches 451; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2790 CCCCCTCAGGAAGAAGGACTATATTGTACTGTACCTAGGCTTCTGGAGGAAAAC 2849

Db |||||

Db 453 CCCCCTCAGGAAGAAGGACTATATTGTACTGTACCTAGGCTTCTGGAGGAAAAC 394

QY 2850 ATGGAATCAGGATTTCTATAGACTGATAGCCCTATCCACAAGGCCATGACTGGGAAAAG 2909

Db |||||

Db 393 ATGGAATCAGGATTTCTATAGACTGATAGCCCTATCCACAGGGCCATGACTGGGAAAAG 334

QY 2910 GTATGGGAGCAGAGGAGAAATTTGGGATTTTAGGTGTCAGCTACGCTACCCCTAACTTTT 2969

Db |||||

Db 333 GTATGGGAGCAGAGGAGAAATTTGGGATTTTAGGTGTCAG-TACGCTACCCCTAACTTTT 275

QY 2970 GGTGGCCTGGGCATGCTCTTGAGGCCCAGACTGTTAAGCAGGCTCTGCTGGCTGTTTAC 3029

Db |||||

Db 274 GGTGGCCTGGGCATGCTCTTGAGGCCCAGACTGTTAAGCAGGCTCTGCTGGCTGTTTAC 215

QY 3030 TCGTCACCACTCTGCACCTGCTGCTTTGAGACTCCATCCAGCCCGCCAGCCACCTG 3089

Db |||||

Db 214 TCGTCACCACTCTGCACCTGCTGCTTTGAGACTCCATCCAGCCCGCCAGCCACCTG 155

QY 3090 CTCCTGAGCCTCCACTATCTCCCTGTGACGGGTGAACCTTCGTGTACTGTCTCGGGTCC 3149

```

Db      154 CTCCTGAGCCTCCACTATCTCCCTGTGACGGGTGAACCTCGTGTACTGTCTCGGGTCC 95
Qy      3150 ATATGTAATTGTGAGCAGGGTTCATCTATTTTAAACACAGATGTTTACAAAAATAAGAT 3209
Db      94 ATATATGAATTGTGAGCAGGGTTCATCTATTTTAAACACAGATGTTTACAAAAATAAGAT 35
Qy      3210 TATTTCACCAACCAAAAAAATAAAAAAATAAAAAA 3243
Db      34 TATTTCACCAACCAAAAAAATAAAAAAATAAAAAA 1

RESULT 12
AAT25457
ID AAT25457 standard; cDNA to mRNA; 320 BP.
XX AC
XX AC
XX AC
XX DT- 17-OCT-1996 (first entry)
XX DE Human gene signature HUMGS07624.
XX KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
XX KW human; cloning; mapping; non-biased library; diagnosis; detection;
XX KW cell typing; abnormal cell function; ss.
XX OS Homo sapiens.
XX PN W09514772-A1.
XX PD 01-JUN-1995.
XX PF 11-NOV-1994; 94WO-JP01916.
XX PR 12-NOV-1993; 93JP-0355504.
XX PA (MATS/) MATSUBARA K.
XX PA (OKUB/) OKUBO K.
XX PI Matsubara K, Okubo K;
XX WPI; 1995-206931/27.
XX DR
XX DR
XX PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
XX PT for diagnosis of abnormal cell function, by preparing cDNA that
XX PT reflects relative abundance of corresp. mRNA in specific human
XX PT tissues
XX PS Claim 1; Page 1835; 2245pp; Japanese.
XX CC A single-stranded DNA (or its complementary strand or the corresp.
XX CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
XX CC given in AAT19001-726837 and which is able to hybridise to part of
XX CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
XX CC sequences were obtained from 3'-directed cDNA libraries prepared
XX CC from various human tissues; synthesis of cDNA was initiated from the
XX CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
XX CC untranslated sequence is unique to a particular mRNA species, almost
XX CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
XX CC is constructed so as to reflect accurately the relative abundance of
XX CC different mRNAs in the particular tissue from which it was derived.
XX CC The appearance frequency of a given GS in a cDNA library can be
XX CC determined (esp. using primers and probes derived from the GS
XX CC sequences) as a means of diagnosing abnormal cell function or for
XX CC recognising different cell types.
XX SQ Sequence 320 BP; 74 A; 75 C; 86 G; 81 T; 4 other;

Query Match 9.6%; Score 313.4; DB 16; Length 320;
Best Local Similarity 98.7%; Pred. No. 4.1e-66;
Matches 314; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      2752 GATCAGCAGGGTGCCATTCTATGTTCTCTCTAGCCCCCTCAGGAAGAGGACTA 2811

```

```

Db      1 GATCAGCAGGGTGCCATTNATTGTTCTTCTCTAGCCCCCTCAGGAAGAGGACTA 60
Qy      2812 TATTTGTACTGTACTCCCTAGGGGTTCTGGAAGGGAACATGGAATCAGGATTCATATAGAC 2871
Db      61 TATTTGTACTGTACTCCCTAGGGGTTCTGGAAGGGAACATGGAATCAGGATTCATATAGAC 120
Qy      2872 TGATAGCCCTATCCACAAGGGCCATGACTGGGAAAAGGATATGGGAGCAGAGGAATT 2931
Db      121 TGATAGCCCTATCCACAAGGGCCATGACTGGGAAAAGGATATGGGAGCAGAGGAATT 180
Qy      2932 GGGATTTTAGGTGCAGCTACGCTACCCCTAAACTTTTGGTGGCCTGGGGCATCTCTTGA 2991
Db      181 GGGATTTNAGGGTGAGCTAGCTACCCCTAAACTTTTGGTGGCCTGGGGTATCTCTTGA 240
Qy      2992 GGCCAGACTGTTAAGCAGGCTCTGCTGGGCTGTTTACTGCTACCCACCTCTGCACCTGC 3051
Db      241 GGCCAGACTGTTAAGCAGGCTCTGCTGGGCTGTTTACTGCTACCCACCTCTGCACCTGC 300
Qy      3052 TGCTTGAGACTCCATCC 3069
Db      301 TGTCTTGAGACTCCATCC 318

RESULT 13
ABL38341
ID ABL38341 standard; cDNA; 313 BP.
XX AC
XX AC
XX AC
XX DT 08-APR-2002 (first entry)
XX DE Human colon tumour antigen polynucleotide SEQ ID NO:1930.
XX KW Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
XX KW colon tumour metastatic antigen; diagnosis; gene; ss.
XX OS Homo sapiens.
XX PN W0200196388-A2.
XX PD 20-DEC-2001.
XX PF 08-JUN-2001; 2001WO-US18557.
XX PR 09-JUN-2000; 2000US-210899p.
XX PR 20-FEB-2001; 2001US-270216p.
XX PS (CORI-) CORIXA CORP.
XX PI Jiang Y, Harlocker SL, Secrlist H;
XX WPI; 2002-114514/15.
XX PT Novel isolated colon tumor polynucleotide differentially expressed in
XX PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
XX PT useful for inhibiting development of cancer in patient -
XX PS Claim 1; SEQ ID 1930; 105pp; English.
XX CC ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
XX CC which were isolated from human colon tumour and colon metastatic tumour
XX CC cDNA libraries. (I) have cytostatic activity and can be used in vaccine
XX CC production. (I) can be used for stimulating and/or expanding T cells
XX CC specific for a tumour protein on contact with the T cells. They are also
XX CC useful for inhibiting the development of cancer in a patient. (I) can be
XX CC used as probes or primers for nucleic acid hybridisation, for preparing
XX CC mutant species primers, or primers for use in genetic constructions. (I)
XX CC can be used in the diagnosis of a colon tumour.
XX SQ Sequence 313 BP; 68 A; 85 C; 85 G; 73 T; 2 other;

Query Match 9.4%; Score 307; DB 24; Length 313;

```

[illegible]

THIS PAGE BLANK (USPTO)

	1	2	3	4	5	6
2690.4	82.6	2700	11	BC014511	BC014511	Homo sapi
947.8	29.1	1081	13	BM551913	BM551913	AGENCOURT
856	26.3	957	14	BQ937298	BQ937298	AGENCOURT
840.2	25.8	908	14	BQ716774	BQ716774	AGENCOURT
810.6	24.9	901	13	BQ220195	BQ220195	AGENCOURT
801.6	24.6	924	13	B1118034	B1118034	602867301


```

QY 2281 CAGCCAAACAAAGAACTGGGTGTGAGTATTATCACTAAGAACCAAAATCCAGGCA 2340
Db 1741 CAGCCAAACAAAGAACTGGGTGTGAGTATTATCACTAAGAACCAAAATCCAGGCA 1800
QY 2341 CTCTATGTGAAGGATAAGAACTCACTTCCTTACTCCTCCAAAGAAAGTGGGGAAGA 2400
Db 1801 CTCTATGTGAAGGATAAGAACTCACTTCCTTACTCCTCCAAAGAAAGTGGGGAAGA 1860
QY 2401 ACCATCAAACTTTCCTCTCTGACTTACCAACACGAGGAAACACAGCAGGAGGCTGCTCA 2460
Db 1861 ACCATCAAACTTTCCTCTCTGACTTACCAACACGAGGAAACACAGCAGGAGGCTGCTCA 1920
QY 2461 GGACTTAGGACAGGATATAGCTTAGTGTGGAAGCAAGAGAGCAGGAGTGTAA 2520
Db 1921 GGACTTAGGACAGGATATAGCTTAGTGTGGAAGCAAGAGAGCAGGAGTGTAA 1980
QY 2521 ATCACTGGCTTAATGAGAAAGGAGACAGCTAAGTCTAGGATGAAGCTGTCACTAGGCTGG 2580
Db 1981 ATCACTGGCTTAATGAGAAAGGAGACAGCTAAGTCTAGGATGAAGCTGTCACTAGGCTGG 2040
QY 2581 AGTTGCTTCCCTTGAAGATGGGACTCCCTTGGGTATCAAGACCTATGCCACATCACACTGGG 2640
Db 2041 AGTTGCTTCCCTTGAAGATGGGACTCCCTTGGGTATCAAGACCTATGCCACATCACACTGGG 2100
QY 2641 GCTAGGGAAGTAGTGATGCCGCCCTCAAGTCTGCTTCCAGCCAGGAGCTTGAAGATT 2700
Db 2101 GCTAGGGAAGTAGTGATGCCGCCCTCAAGTCTGCTTCCAGCCAGGAGCTTGAAGATT 2160
QY 2701 ATATTGGGCACTGGCTCAATCTGTGGACCAATTTTCAGCTTTCCTGAAGATCAGGCA 2760
Db 2161 ATATTGGGCACTGGCTCAATCTGTGGACCAATTTTCAGCTTTCCTGAAGATCAGGCA 2220
QY 2761 GGSTGGCAATTCATTTCTCTCTCTAGCCCTCAGGAAAGAGACTATATTGTAC 2820
Db 2221 GGSTGGCAATTCATTTCTCTCTCTAGCCCTCAGGAAAGAGACTATATTGTAC 2280
QY 2821 TGTACCTTAGGGTCTCGGAAGGAAACATGGAATCAGATTCTATAGACTATAGGCC 2880
Db 2281 TGTACCTTAGGGTCTCGGAAGGAAACATGGAATCAGATTCTATAGACTATAGGCC 2340
QY 2881 CTATCCACAGGGCCATGACTGGGAAAGGATGGGAGCAGAGGAGAAATGGGATTATA 2940
Db 2341 CTATCCACAGGGCCATGACTGGGAAAGGATGGGAGCAGAGGAGAAATGGGATTATA 2400
QY 2941 GGSTGACGCTACGCTACCCCTAACTTTTGGTGGCTGGGCGATGCTTGAAGCCAGAC 3000
Db 2401 GGSTGACGCTACGCTACCCCTAACTTTTGGTGGCTGGGCGATGCTTGAAGCCAGAC 2460
QY 3001 TGTAAAGCAGGCTCTGCTGGCTGTTTACTCTGTCACACCTCTGCACCTCTGCTTGAG 3060
Db 2461 TGTAAAGCAGGCTCTGCTGGCTGTTTACTCTGTCACACCTCTGCACCTCTGCTTGAG 2520
QY 3061 ACTCCATCCAGCCCGCAGGACGACCTGCTCTGAGCCCTCCACTATCTCCCTGTGACGG 3120
Db 2521 ACTCCATCCAGCCCGCAGGACGACCTGCTCTGAGCCCTCCACTATCTCCCTGTGACGG 2580
QY 3121 GTGAACCTGCTGTACTGTCTCGGCTCCATATATGAATTTGTGAGCAGGGTTTCATCTATT 3180
Db 2581 GTGAACCTGCTGTACTGTCTCGGCTCCATATATGAATTTGTGAGCAGGGTTTCATCTATT 2640
QY 3181 TTAACACAGATGTTTACAAAATAAGATTATTTCAACCAACCAAAAAA 3240
Db 2641 TTAACACAGATGTTTACAAAATAAGATTATTTCAACCAACCAAAAAA 2700

```

RESULT 2

```

BM551913 LOCUS 1081 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT_6575472 NIH_MGC_98 Homo sapiens cdna clone IMAGE:5479433
5', mRNA sequence.
ACCESSION BM551913
VERSION BM551913.1 GI:18789340
KEYWORDS EST.

```

SOURCE

ORGANISM

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: rgs@nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLC2000 row: k column: 18

High quality sequence stop: 662.

Location/Qualifiers

1. 1081

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5479433"

/lab_host="NIH_MGC_98"

/tissue_type="astrocytoma grade IV, cell line"

/note="Organ: brain; Vector: pOTB7; Site: 1: XhoI; Site 2:

EcORI; CDNA made by oligo-dr priming. Directionally

cloned into EcORI/XhoI sites using the following 5'

adaptor: GGCAGGAG(G). Library constructed by Ling Hong

in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-CDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH_MGC Library."

BASE COUNT 260 a 295 c 295 g 229 t

ORIGIN

Query Match 29.1%; Score 947.8; DB 13; Length 1081;

Best Local Similarity 95.4%; Pred. No. 1.4e-153;

Matches 1029; Conservative 0; Mismatches 44; Indels 6; Gaps 5;

QY 1178 GCTATCCAAAGGGATGACCCAGGCTCGTGGGAGGTTCCACCTTGGATCCAGGAAGA 1237

Db 1 GCTATCCAAAGGGATGACCCAGGCTCGTGGGAGGTTCCACCTTGGATCCAGGAAGA 60

QY 1238 CCCTCCACCTGCTGCTGCTGGTGGCCAAAGGCTACAGGCTGCTTCTTCTTCCCTCC 1297

Db 61 CCCTCCACCTGCTGCTGCTGGTGGCCAAAGGCTACAGGCTGCTTCTTCTTCCCTCC 120

QY 1298 CCCCCACTGTCCCTCATGTGCCATGGCCCTGCCCTCCCTCCAGTGACCTCGGAAAGTGGAGCA 1357

Db 121 CCCCCACTGTCCCTCATGTGCCATGGCCCTGCCCTCCCTCCAGTGACCTCGGAAAGTGGAGCA 180

QY 1358 TCAGGTAGGAGGAAACACACACCGGGGAGTCTTCAGGCTTGGGGTGCCTTACCTCTA 1417

Db 181 TCAGGTAGGAGGAAACACACACCGGGGAGTCTTCAGGCTTGGGGTGCCTTACCTCTA 240

QY 1418 CCATTCCCGGACGACAGCTTTCCTTGGCTGCTGCTGCCCTGCCCTTCTTGGGGACT 1477

Db 241 CCATTCCCGGACGACAGCTTTCCTTGGCTGCTGCTGCCCTGCCCTTCTTGGGGACT 300

QY 1478 GAGCTCAGAGCGAGGCTTTCAGAGAGGAAACAAATAGGGGTGGCAGGATATAAAG 1537

Db 301 GAGCTCAGAGCGAGGCTTTCAGAGAGGAAACAAATAGGGGTGGCAGGATATAAAG 360

QY 1538 TCACCTCCATTCCTACCTCCCATGACATGAACAATTCCTTCTCCACTGCTCCCA 1597

Db 361 TCACCTCCATTCCTACCTCCCATGACATGAACAATTCCTTCTCCACTGCTCCCA 420

QY 1598 AATTAAAGATGCGACCAAGGCTTGGGTCTTCCAGGGGCGAGGAGCCCTGGGGTC 1657

Db 421 AATTAAAGATGCGACCAAGGCTTGGGTCTTCCAGGGGCGAGGAGCCCTGGGGTC 480

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13331 row: e column: 07
High quality sequence stop: 606.

FEATURES
Source 1..901
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6061422"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: PCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 251 a 203 c 242 g 205 t

ORIGIN

Query Match 24.9%; Score 810.6; DB 14; Length 901;
Best Local Similarity 98.7%; Pred. No. 6.2e-130;
Matches 838; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

QY 2160 CTTTAAAGACGGAACAAAGTAATTTACCACTTCTACTGGGTTCCTGCCACCGTCCCAAG 2219
DB 1 CTTTAAAGACGGAACAAAGTAATTTACCACTTCTACTGGGTTCCTGCCACCGTCCCAAG 60

QY 2220 GTGGCGAGGCGCTAGGAAGAGGGTCATCTTAAGCCACACATAGTCGCACCTGCGTGGCT 2279
DB 61 GTGGCGAGGCGCTAGGAAGAGGGTCATCTTAAGCCACACATAGTCGCACCTGCGTGGCT 120

QY 2280 GCAGCCAAACAAAGACACTGGGTGTGAGTATTCATCACTAAGAACCAAAATCCAGGC 2339
DB 121 GCAGCCAAACAAAGACACTGGGTGTGAGTATTCATCACTAAGAACCAAAATCCAGGC 180

QY 2340 ACTCATATGTGAAGATAAGAACTCCTTCTTACTCTCCAAAAGAAAGTGGGAAG 2399
DB 181 ACTCATATGTGAAGATAAGAACTCCTTCTTACTCTCCAAAAGAAAGTGGGAAG 240

QY 2400 ACCATCAACCTTTCTCTGACITACCAACACGAGAAACAGCAGGAGGGTGGCTC 2459
DB 241 ACCATCAACCTTTCTCTGACITACCAACACGAGAAACAGCAGGAGGGTGGCTC 300

QY 2460 AGGACTTAGGACAGGGTATAGCTTAGATGGTGGAAAGCAAGGAGCAGGAAGTTGTA 2519
DB 301 AGGACTTAGGACAGGGTATAGCTTAGATGGTGGAAAGCAAGGAGCAGGAAGTTGTA 360

QY 2520 AATCACTGGCTAATGAAAGAGACAGTCACTTCTAGGATGAAGCTGTAGTGGCTG 2579
DB 361 AATCACTGGCTAATGAAAGAGACAGTCACTTCTAGGATGAAGCTGTAGTGGCTG 420

QY 2580 GAGTTCCTTCTGAGATGGGACTCCTTGGGTATCAACACCTATGCCACATCACTGG 2639
DB 421 GAGTTCCTTCTGAGATGGGACTCCTTGGGTATCAACACCTATGCCACATCACTGG 480

QY 2640 GGCTAGGGAAGTAGGTGATGCCAGCCCTCAAGTGTGTCTTTCAGCCAGGACTTGAGAAGT 2699
DB 481 GGCTAGGGAAGTAGGTGATGCCAGCCCTCAAGTGTGTCTTTCAGCCAGGACTTGAGAAGT 540

QY 2700 TATATTGGCAGTGGCTCCAAATCTGTGGACCACTATTTTCAGCTTTCCCTTGAAGATCAGGC 2759
DB 541 TATATTGGCAGTGGCTCCAAATCTGTGGACCACTATTTTCAGCTTTCCCTTGAAGATCAGGC 600

QY 2760 AGGGTCCCATTCATCTCTTCTCTAGCCCTCCTCAGGAAGAGGACTATATTGTA 2819
DB 601 AGGGTCCCATTCATCTCTTCTCTAGCCCTCCTCAGGAAGAGGACTATATTGTA 660

QY 2820 CTGTACCCCTAGGGTCTTGGGAAGGGAACATGGAATCAGGATCTTATAGACTGATAGGC 2879
DB 661 CTGTACCCCTAGGGTCTTGGGAAGGGAACATGGAATCAGGATCTTATAGACTGATAGGC 720

QY 2880 CCTATCCACAAGGCGCATGACTGGGAAAAGCTATGGGAGCAG-AAGGAGAATTCGGATTT 2938
DB 721 CCTATCCACAAGGCGCATGACTGGGAAAAGCTATGGGAGCAGAAAGGAATTCGGATTT 780

QY 2939 TAGGTGCAGCTACGCTCACCTAAACTTTTGGTGGCTT-GGGGCATCTCTTGAGGCCCA 2997
DB 781 TAGGTGCAGCTACGCTCACCTAAACTTTTGGGCGCTTGGGCGATCTCTTGAGGCCCA 840

QY 2998 GACTCTTAA 3006
DB 841 GACTGGTTA 849

RESULT 6
BI118034
LOCUS
DEFINITION
ACCESSION BI118034
VERSION BI118034.1 GI:14568935
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 924)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM1825 row: b column: 02
High quality sequence stop: 893.

FEATURES
Source 1..924
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5015713"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 260 a 225 c 244 g 195 t

ORIGIN

Query Match 24.6%; Score 801.6; DB 13; Length 924;
Best Local Similarity 93.7%; Pred. No. 2.2e-128;
Matches 868; Conservative 0; Mismatches 54; Indels 4; Gaps 3;

QY 1504 AGGAACAAAATGAGGGTGGCAGGAGTAAAGATCAATCTCTTACTCTCCATGC 1563
DB 2 AGGAACAAAATGAGGGTGGCAGGAGTAAAGATCAATCTCTTACTCTCCATGC 61

QY 1564 AGCATGAACACAATTTCTCTCCACCTGGCTCCCAAAATTAAGATGTGGACCAAGGCCTG 1623

```
Db 62 AGCATGACAAATTTCTCTCCACTGGCTCCCAATTTAAAGATGGCAAGGCGTG 121
Qy 1624 TGGTACTCCAGGGGCAAGAGAGCCCTGGGGTCACTGACACTCTCAGGCCCAACCATGCA 1683
Db 122 TGGTACTCCAGGGGCAAGAGAGCCCTGGGGTCACTGACACTCTCAGGCCCAACCATGCA 181
Qy 1684 CTCACAAAGGGGAGCAATTTGGAAATGAAGACTAGCTCTCTATATCAGTTTAAGAGCA 1743
Db 182 CTCACAAAGGGGAGCAATTTGGAAATGAAGACTAGCTCTCTATATCAGTTTAAGAGCA 241
Qy 1744 AGGAGAGCTGGCCAGGAGCAGCACTTTGACACAGAGGGGAATGTAAGCAACAGCAGG 1803
Db 242 AGGAGAGCTGGCCA-GGACAGCACTTTGACACAGAGGGGAATGTAAGCAACAGCAGG 300
Qy 1804 CCTCTAGGCCCTTCTCCATTTCTTAGTAAAGAGCAATTTCTCAGACTCCCAAGGC 1863
Db 301 CCTCTAGGCCCTTCTCCATTTCTTAGTAAAGAGCAATTTCTCAGACTCCCAAGGC 360
Qy 1864 GGAGACTAGCCCTAGCCCTCAGCAACCAAGGTTCTCTGGGACCCAAAGTTTATGGAG 1923
Db 361 GGAGACTAGCCCTAGCCCTCAGCAACCAAGGTTCTCTGGGACCCAAAGTTTATGGAG 420
Qy 1924 AAGGCAAGACTTCATGGGAAGAGAGAGGAAGGCCCTGGGTAGAAACGCTTGTGCTG 1983
Db 421 AAGGCAAGACTTCATGGGAAGAGAGAGGAAGGCCCTGGGTAGAAACGCTTGTGCTG 480
Qy 1984 TTCTCTTTGGCTTTAAGACAAAGCGCTCATCTTTGCCCTCTACCTCTCTGATAGCTTGAG 2043
Db 481 TTCTCTTTGGCTTTAAGACAAAGCGCTCATCTTTGCCCTCTACCTCTCTGATAGCTTGAG 540
Qy 2044 GGTTCGCAACCACTGTGGCTACAGTGGAGGAAGAGGACTCTCTCTCCAGAGTGC 2103
Db 541 GGTTCGCAACCACTGTGGCTACAGTGGAGGAAGAGGACTCTCTCTCCAGAGTGC 600
Qy 2104 TATGTTGAGGAAGTTCTTTAAACCCATATATGGCCCAAGAGTAGCTCTGAGAGCCCTTT 2163
Db 601 TATGTTGAGGAAGTTCTTTAAACCC--ATATGGCCCAAGAGTAGCTCTGAGAGCCCTTT 658
Qy 2164 AAAGCAGGAACAAGTAATTTACCAAGTTCTACTG--GGTTCTCTGCCCAAGAGTGC 2222
Db 659 AAAGCAGGAACAAGTAATTTACCAAGTTCTACTGAGGTTCCCTGGCCACCGTCCCAAGTG 718
Qy 2223 GGCAGGCTAGGAAGAGGTCATCTTTAAGCCACATAGCTAGCTAGCTGGTGGCTGCA 2282
Db 719 GGCAGGCTAGGAAGAGGTCATCTTTAAGCCACATAGCTAGCTAGCTGGTGGCTGCA 778
Qy 2283 GCCAAACAAAGAACTGGGTGTGAGTATTCATCAACTAAGAAACCAAAATCCAGGCACT 2342
Db 779 GCCAAACAGAACTGGGTGTGAGTATTCATCAAGTAAGAAACCAAAATCCAGGCACT 838
Qy 2343 CATATGTGAAGGATGAAGAACTCTCTCTTACTCTCTCCAAAGAAAGTGGGGAAGAAC 2402
Db 839 CATATGTGAAGGATGAAGAACTCTCTCTTATCTCTCTCCAAAGAAAGTGGGGAAGAAC 898
Qy 2403 CATCAACCTTTCTCTCTGACTTACC 2428
Db 899 ATTAACCTTACTCTGATTACAAAGC 924
```

RESULT 7

```
BQ277739
LOCUS BQ277739 1046 bp mRNA linear EST 07-MAY-2002
DEFINITION AGENCOURT_7047903 NIH_MGC_109 Homo sapiens cDNA clone IMAGE:5804597
5', mRNA sequence.
ACCESSION BQ277739
VERSION BQ277739.1 GI:20487947
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1046)
```

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LCM2042 row: h column: 06
High quality sequence stop: 570.

FEATURES
source

Location/Qualifiers
1..1046
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5804597"
/lab_host="NIH_MGC_109"
/tissue_type="teratocarcinoma, cell line"
/note="Organ: ovary; Vector: pOTB7; Site_1: EcoRI; Site_2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 297 a 250 c 291 g 208 t
ORIGIN

Query Match 24.5%; Score 799.2; DB 14; Length 1046;
Best Local Similarity 96.9%; Pred. No. 5.4e-128;
Matches 847; Conservative 0; Mismatches 23; Indels 4; Gaps 3;
Qy 1547 TTCCTACCTCCCATGAGCATGAACAAATTTCTCTCCACTGGCTCCCAATTTAAAG 1606
Db 1 TCTCTACTCTCCCATGAGCATGAACAAATTTCTCTCCACTGGCTCCCAATTTAAAG 60
Qy 1607 ATGTGGACCAAGGCTGTGGTACTCCAGGGGCAAGAGAGCCCTGGGTGAGTCACT 1666
Db 61 ATGTGGACCAAGGCTGTGGTACTCCAGGGGCAAGAGAGCCCTGGGTGAGTCACT 120
Qy 1667 GTCAGGCCAACCATGCACTCCACAAAGGGGAGCATTTGGAATGAAGACTAGCTCTAT 1726
Db 121 GTCAGGCCAACCATGCACTCCACAAAGGGGAGCATTTGGAATGAAGACTAGCTCTAT 180
Qy 1727 GTATCAGTTTAAGAGCAAGGAGAGCTGGCCAGGACAGCATGTTGCAACAGAGGGA 1786
Db 181 GTATCAGTTTAAGAGCAAGGAGAGCTGGCCAGGACAGCATGTTGCAACAGAGGGA 240
Qy 1787 ATGTAGCAACAGAGGCTCTAGGCCCTCTAGGCCCTCTAGGCCCTCTAGGCCCT 1846
Db 241 ATGTAGCAACAGAGGCTCTAGGCCCTCTAGGCCCTCTAGGCCCTCTAGGCCCT 300
Qy 1847 TCCTCAGACTCCAGGGGAGGACTGAGCCTTAGCCTTCCAGCAACCAAGGTTCTCT 1906
Db 301 TCCTCAGACTCCAGGGGAGGACTGAGCCTTAGCCTTCCAGCAACCAAGGTTCTCT 360
Qy 1907 CCCAAAGTTTATGGGAGAAAGCAAGACTTTCATGGGAGAGAGAGGAGGAGGAGG 1966
Db 361 CCCAAAGTTTATGGGAGAAAGCAAGACTTTCATGGGAGAGAGAGGAGGAGGAGG 420
Qy 1967 AGAAACGGTGTGCTGTCTCTTTGGCTTTAAGACAAAGGCTCATCTTGCCTCTAC 2026
Db 421 ACAACGGTGTGCTGTCTCTTTGGCTTTAAGACAAAGGCTCATCTTGCCTCTAC 480
Qy 2027 CTCCTGATAGCTTGAAGGTTTGGCAACCACTGCTGCTAGGAGGAGGAGGAGGAG 2086
Db 481 CTCCTGATAGCTTGAAGGTTTGGCAACCACTGCTGCTAGGAGGAGGAGGAGGAG 540

QY	2087	TCCTTCTCAGAGTCTATGTTCAGGAAGTTCTTTAACCCCATATAGCCCAAGAGTAG	2144
Db	541	TCCTTCTCAGAGTCTATGTTCAGGAAGTTCTTTAACCCCATATAGCCCAAGAGTAG	600
QY	2147	CTCGTAGGAGGCCCTTTAAAGACGGAAACAAGTAATTTACAGTTCTACTGGGGTTCCTGC	2206
Db	601	CTCCTATGAGGCCCTTTAAAGACGGAAACAAGTAATTTACAGTTCTACTGGGGTTCCTGC	660
QY	2207	CCACCGTCCCAAGTGGCGAGGCCCTTAGGAAGAGGTCTATCTTAAGCCACACATTAGCT	2266
Db	661	CCACCGTCCCAAGTGGCGAGGCCCTTAGGAAGAGGTCTATCTTAAGCCACACATTAGCT	720
QY	2267	GCACTCGTGGCTGCAGCCCAAAACAAGAACTGGGTGTTGAGTATTCATCAACTAAGAAC	2326
Db	721	GCACTCGTGGCTGCAGCCCAAAACAAGAACTGGGTGTTGAGTATTCATCAACTAAGAAC	780
QY	2327	CAAAATCCAGGGCACTCATATGT-GAAGGATAGAACC--TCATCTCCTTACTCCTCCAA	2383
Db	781	CAAAATCCGGGGCACTCATATGTGGAAGGATAGAACCCTTCTTCTTACTCCTCCAA	840
QY	2384	AAAGAACTGGGG-AAAGAACCATCAACCTTCC	2416
Db	841	AAAAAGGGGGAAAAAGAACCATCGAACTTTTC	874
RESULT 8			
BM476862			
LOCUS	BM476862	1027 bp	mRNA linear EST 05-FEB-2002
DEFINITION	AGENCOURT_6481739 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5555414		
5', mRNA sequence.			
ACCESSION	BM476862		
VERSION	BM476862.1 GI:18525904		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 1027)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM12275 row: i column: 15 High quality sequence stop: 607.		
FEATURES			
source			
1..1027			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="IMAGE:5555414"			
/clone_lib="NIH_MGC_71"			
/tissue_type="leiomyosarcoma"			
/lab_host="DH10B (phage-resistant)"			
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.			
Average insert size 2.1 kb.			
BASE COUNT	260 a	272 c	283 g
ORIGIN	211 t	1 others	
Query Match			
Best Local Similarity 24.0%; Score 781.2; DB 13; Length 1027;			
Matches: 823; Conservative 0; Mismatches 14; Indels 4; Gaps 3;			
QY	1354	AGCATCAGGTAGGAGGAAACAGCAACCGGGAGTCTCTGAGCCTGGGGTGCCTTACC	1413
Db	2	AGCATCAGGTAGGAGGAAACAGCAACCGGGAGTCTCTGAGCCTGGGGTGCCTTACC	61

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLNL at:
http://image.llnl.gov
Plate: L1AM10600 row: a column: 14
High quality sequence stop: 768.
Location/Qualifiers
1..878

FEATURES
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4748437"
/clone_lib="NCI_CGAP_Skn4"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI CGAP Library."
Technologies. Note: this is a NCI CGAP Library."
BASE COUNT 249 a 200 c 237 g 191 t 1 others
ORIGIN
Query Match 23.6%; Score 767.4; DB 12; Length 878;
Best Local Similarity 97.4%; Pred. No. 1.7e-122;
Matches 833; Conservative 0; Mismatches 16; Indels 6; Gaps 5;
QY 1822 CCATTTCTTAGTAAAGAGAGATTTCTCAGACTCCAGCGGAGGAGCTAGCGCTAGCC 1881
DB 1 CCATTTCTTAGTAAAGAGAGATTTCTCAGACTCCAGCGGAGGAGCTAGCGCTAGCC 60
QY 1882 TTGAGCAACCAAGGTTCTCTGGGACCCAAAGTTATGGGAGAGGCGCAAGACTCATG 1941
DB 61 TTGAGCAACCAAGGTTCTCTGGGACCCAAAGTTATGGGAGAGGCGCAAGACTCATG 120
QY 1942 GGAAGAGAGAGAGAGGCGCTGGGTAGAAACCGCTTGGTGTCTTTGGCTTTTAAAG 2001
DB 121 GGAAGAGAGAGAGAGGCGCTGGGTAGAAACCGCTTGGTGTCTTTGGCTTTTAAAG 180
QY 2002 ACAAAGGCTCATCTTGGCTCTACCTCTGTATAGGCTTGGGTTTGCCAAACCACTG 2061
DB 181 ACAAAGGCTCATCTTGGCTCTACCTCTGTATAGGCTTGGGTTTGCCAAACCACTG 240
QY 2062 TGGCTACAGGTGGAGGAGAGAGACTCTTCTCCAGAGTGTATGTTCAAGAGTTTCT 2121
DB 241 TGGCTACAGGTGGAGGAGAGAGACTCTTCTCCAGAGTGTATGTTCAAGAGTTTCT 300
QY 2122 TTAACCCCATATGCCCCAAGAGTAGTCTGTAGGAGGCGCTTTAAAGACGGAACAAGTAAT 2181
DB 301 TTAACCCCATATGCCCCAAGAGTAGTCTGTAGGAGGCGCTTTAAAGACGGAACAAGTAAT 360
QY 2182 TTACCACTTCTACTGGGGTTCCTCCACCGTCCCAAGTGGGCGAGGCTTAGAAGAGG 2241
DB 361 TTACCACTTCTACTGGGGTTCCTCCACCGTCCCAAGTGGGCGAGGCTTAGAAGAGG 420
QY 2242 GTCACTTCTTAAGCCACACATAGCTGTGACCTGCGTGGCTGAGCGCAAAACAAAGACTGGG 2301
DB 421 GTCACTTCTTAAGCCACACATAGCTGTGACCTGCGTGGCTGAGCGCAAAACAAAGACTGGG 480
QY 2302 TGTGAGTATTCATCACTAAGAACCAAAATCCAGGCGACTCATATGTGAAGGATAAGAA 2361
DB 481 TGTGAGTATTCATCACTAAGAACCAAAATCCAGGCGACTCATATGTGAAGGATAAGAA 540
QY 2362 CCTCACTTCTTACTCTCCAAAGAGAGTGGGGAAGACCATCAACCTTTCCTCTG 2421
DB 541 CCTCACTTCTTACTCTCCAAAGAGAGTGGGGAAGACCATCAACCTTTCCTCTG 600
QY 2422 ACTTACCAACCAAGGAGAGAGGAGGCTGAGGAGGCTTAGGAGCAGGCTATAG 2481

Db 601 ACTTACCAACCAAGGAGAGAGAGAGAGGCTGAGGCTAGGAGCAGGCTATAG 660
QY 2482 CTTAGATGTTGGAAGCAAGAGAGAGAGAGGCTGAGGCTATAGGAGAGAG 2541
Db 661 CTTAGATGTTGGAAGCAAGAGAGAGAGAGGCTGAGGCTATAGGAGAGAG 720
QY 2542 GAGACAGCTAATCTCTAGGATGAGCTGTGACTAGG-CTGAGTTGCTTCTTGA-AGATG 2599
Db 721 GAGA-AGCTAATCTAGGATGAGCTGTGACTAGGCTGAGGCTGAGTTGCTTCTTGAACAAC 779
QY 2600 GGACTCTCTGGTATCAAGACCTA-TGCCACATCACACTG--GGCTAGGAGAGTGTG 2656
Db 780 GGACTCTCTGGTATCAGACCTATGCCACATCACACTGAGCTAGGAGAGAGTGTG 839
QY 2657 ATGCCAGCCCTCAAG 2671
Db 840 ATGCAAGGCTCAAG 854
RESULT 10
BM977746/c
LOCUS
DEFINITION
UI-CF-EN1-aef-h-15-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
UI-CF-EN1-aef-h-15-0-UI 3', mRNA sequence.
ACCESSION
BM977746
VERSION
BM977746.1 GI:19596476
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 767)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
9704477
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers
1..767
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-CF-EN1-aef-h-15-0-UI"
/clone_lib="UI-CF-EN1"
/tissue_type="Primary Lung Cystic Fibrosis Epithelial
Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-CF-EN1 is a normalized cDNA library containing the
following tissue(s): Primary Lung Cystic Fibrosis
Epithelial Cells. The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pT73-Pac vector. The
oligonucleotide used to prime the synthesis of

Qy	2655	TGATGCCAGCCCTCAAGTC	TGCTTCCAGCCAGGAGCTTGAGAA	GTATATATTGGGCAGTGG	2714
Db	584	TGATGCCAGCCCTCAAGTC	TGCTTCCAGCCAGGAGCTTGAGAA	GTATATATTGGGCAGTGG	525
Qy	2715	CTCCAACTCTGTGACCAGTAT	TTTCAGCTTTCCCTGAAGATCAGG	CAGGCTGCCATTCATTT	2774
Db	524	CTCCAACTCTGTGACCAGTAT	TTTCAGCTTTCCCTGAAGATCAGG	CAGGCTGCCATTCATTT	465
Qy	2775	GTCTTTCTCTCCTAGCCCTC	AGGAAGAAGACATATTTCTACT	GTACCCCTAGGGGT	2834
Db	464	GTCTTTCTCTCCTAGCCCTC	AGGAAGAAGACATATTTCTACT	GTACCCCTAGGGGT	405
Qy	2835	TCGTGGAAGGAAACATGGA	ATCTATAGACTGATAGCCCTAT	TCCACAAGGC	2894
Db	404	TCGTGGAAGGAAACATGGA	ATCTATAGACTGATAGCCCTAT	TCCACAAGGC	345
Qy	2895	CATGACTGGGAAAGGTATG	GGAGCAGAAATGGGATTTTAGG	TGTCAGCTACGC	2954
Db	344	CATGACTGGGAAAGGTATG	GGAGCAGAAATGGGATTTTAGG	TGTCAGCTACGC	285
Qy	2955	TCACCTTAAACTTTTGGT	TGGCCCTGGGCGATGCTTGAG	GCCAGACTGTTTAAAGCAGSCTC	3014
Db	284	TCACCTTAAACTTTTGGT	TGGCCCTGGGCGATGCTTGAG	GCCAGACTGTTTAAAGCAGSCTC	225
Qy	3015	TGCTGGCCCTGTTTACTG	TGCACCACTCTGCACCTGCTG	TCTTGAGACTCCATCCAGCCC	3074
Db	224	TGCTGGCCCTGTTTACTG	TGCACCACTCTGCACCTGCTG	TCTTGAGACTCCATCCAGCCC	165
Qy	3075	CAGGCAGCCACCTGCTCCT	GAGCCCTCCACTATCTCCCTG	TGAGCGGTGAACCTTCGTGTA	3134
Db	164	CAGGCAGCCACCTGCTCCT	GAGCCCTCCACTATCTCCCTG	TGAGCGGTGAACCTTCGTGTA	105
Qy	3135	CTGTGCTCTCGGGTCCAT	ATATGAATTTGTAGCAGGGTT	CTATCTATTTTAAACACAGATGT	3194
Db	104	CTGTGCTCTCGGGTCCAT	ATATGAATTTGTAGCAGGGTT	CTATCTATTTTAAACACAGATGT	45
Qy	3195	TTTCAAAATTAAGATTA	TTTCAACACCAACCAAAAAA	AAAAA	3238
Db	44	TTTCAAAATTAAGATTA	TTTCAACACCAACCAAAAAA	AAAAA	1

RESULT 12
 BE799964
 LOCUS
 DEFINITION
 BE799964
 mRNA sequence.
 ACCESSION
 BE799964
 VERSION
 BE799964.1
 KEYWORDS
 GI:10221162
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 990)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTF
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCMT95 row: i column: 15
 High quality sequence start: 22
 High quality sequence stop: 808.
 Location/Qualifiers
 1..990
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 FEATURES
 source

		/clone="IMAGE:3942230"			
		/clone_lib="NIH_MGC_7"			
		/tissue_type="small cell carcinoma"			
		/cell_line="MGC3"			
		/lab_host="DH10B (phage-resistant)"			
		/note="Organ: lung; Vector: pOTB7; Site_1: xhoi; Site_2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/xhoi sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."			
BASE COUNT	229 a	273 c	281 g	206 t	1 others
ORIGIN					
Query Match . 22.9%; Score 746; DB 12; Length 990;					
Best Local Similarity 95.98; Pred. No. 7.9e-119;					
Matches 873; Conservative 0; Mismatches 25; Indels 12; Gaps 10;					
QY	1256	GGGTGGGCCAAAGGCTACAGGGTGCTTCTTCCTCTTCCCCACCCCCACCTGTCCCTCATG	1315		
DB	3	GGCTGGTGTCATCGGCCGAGGCGTCTTCTCTTCCCCACCCCCACCTGTCCCTCATG	62		
QY	1316	TGCATGGGCGCTGCCCTCCCCAGTACACTCGGAAGTGGAGCATCGAGGTAGGAGGGAAC	1375		
DB	63	TGCCATGGGCTGCCTCCCGAGTGACCTCGGAAGTGGAGCATCGAGGTAGGAGGGAAC	122		
QY	1376	AGCAACGGGGAGTCTCTCGAGCTGGGGCTGCCCTACCTCTACCCATTCCCCGACCAG	1435		
DB	123	GGCAACC - AGGAGTCTCGAGCTGGGGCTGCCCTACCTCTACCCATTCCCCGACCAG	181		
QY	1436	CTTTGGCCCTTGGCTGCCCGCGCTGCCTCTTTGGGAACTGAGCTCAGGCGAGGTGC	1495		
DB	182	CTTTGGCCCTTGGCTGCCCGCGCTGCCTCTTTGGGAACTGAGCTCAGGCGAGGTGC	241		
QY	1496	TTCAGAAAGGAACAAATATGAGGGTGGCAGGATATAAAGTACCTCATTTCTTACC	1555		
DB	242	TTCAGAAAGGAACAAATATGAGGGTGGCAGGATATAAAGTACCTCATTTCTTACC	301		
QY	1556	TCCATGCAGCATGAACACAATTTCTCTCACCTGGCTCCCAATTTAAAGATGTGGACC	1615		
DB	302	TCCATGCAGCATGAACACAATTTCTCTCACCTGGCTCCCAATTTAAAGATGTGGACC	361		
QY	1616	AAGSCCTTGGGTACTCCAGGGGCAAGGAGAGCCCTGGGTCACTGACACTGTCCAGGCCA	1675		
DB	362	AAGSCCTTGGGTACTCCAGGGGCAAGGAGAGCCCTGGGTCACTGACACTGTCCAGGCCA	421		
QY	1676	ACCATGCATCTCCACAAAGGGGAGCATTTGGAAATGAAGGACTAGTCTCTATGATCAGGT	1735		
DB	422	ACCATGCATCTCCACAAAGGGGAGCATTTGGAAATGAAGGACTAGTCTCTATGATCAGGT	481		
QY	1736	TAAGACAAGGGAGAGCTGCCCAGGGACACGAGTTTCCACAGCAGAGGGGAATGTAGCAA	1795		
DB	482	TAAGACAAGGGAGAGCTGGGCCA - GGACAGCAGTTTGCACAGCAGAGGGGAATGTAGCAA	540		
QY	1796	CAGCAGGGCTCTCTAGSCCCCATCTCTCCATTTCTTAGTGAAGAGACATTTCTCTCAGAC	1855		
DB	541	CAGCAGGGCTCTCTAGSCCCCATCTCTCCATTTCTTAGTGAAGAGACATTTCTCTCAGAC	600		
QY	1856	TCCCAGCGGAGGACTGAGCTAGCTTTCAGCAACCAAGGTTCTCTCTGGGACCCAA - AGT	1914		
DB	601	TCCCAGCGGAGGACTGAGCTTTCAGCAACCAAGGTTCTCTCTGGGACCCAAAGT	660		
QY	1915	TTATGGGAGA - AGGGCAAGACTTTCATGGGAAGAGAGAGAGGCCCTCGGTAGAAACG	1973		
DB	661	TTATGGGAGACAGGGCAAGACTTTCATGGGAAGAGAGAGAGGCCCTCGGTAGAAACG	720		
QY	1974	CTTGGTCTGTCTCTTTGGCTTTAAGACAAGCGCTCATCTTGGCCCTCTACCTCTCTGA	2033		
DB	721	CTTGGTCTGT - TCTCTTTGGCTTTAAGACAAGCGCTCATCTTGGCCCTCTA - CTCCTGA	778		
QY	2034	TAGGCTTTGAGGGTTTGGCAACC - ACACGTGTGGCT - ACAGCTGGAGGAAAGGAGCTCTCTT	2091		

```

779 TAGGCTTGAGGCTTGGCAACCACTGCTGCTAAACGGTGGAGGCGACGAGCTCCTT 838
2092 CCTCAGAGTCTATGTTTCAGGAAGTTCTTTAAACCCCAATGCGCCCAAGAGTAGCTCGT 2151
839 CCTCAGAGTCTATGTTTCAGGAAGTTCTTTAAACCC---ATTGGCCAGAGTA-CTCGT 894
2152 AGGAGGCCCT 2161
895 AGGAGGCCCT 904

RESULT 13
BM045331
LOCUS 732 bp mRNA linear EST 07-NOV-2001
DEFINITION 60362321F1 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5449420 5',
mRNA sequence.
ACCESSION BM045331
VERSION BM045331.1 GI:16774598
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 732)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1936 row: 1 column: 05
High quality sequence stop: 724.
Location/Qualifiers
1..732
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5449420"
/tissue_lib="NIH_MGC_40"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: prostate; Vector: pOTB7; Site:1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dt priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 118 a 267 c 181 g 166 t
ORIGIN

Query Match 22.2%; Score 722; DB 13; Length 732;
Best Local Similarity 99.3%; Pred. No. 1.2e-114;
Matches 725; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 572 GCTGGAAGCCCTCCGGATGCTGGCCAGCAATGAGTTGGCAGCTGGCAGAGCTGGAGGAGGA 631
Db 1 GCTGGAAGCCCTCCGGATGCTGGCCAGCAATGAGTTGGCAGCTGGCAGAGCTGGAGGAGGA 60
QY 632 GGAGGCAAAACCCGAGGCCCCAGCACTCTTCGAAAACCTGCTGCTGCTGAGTCTCTGTG 691
Db 61 GGAGGCAAAACCCGAGGCCCCAGCACTCTTCGAAAACCTGCTGCTGCTGAGTCTCTGTG 120
QY 692 TGGGGCACCCACAGCACACCCCTCTCCCTCAGGAGGCCCGTGGGCAGACAGGGGAGCC 751
Db 121 TGGGGCACCCACAGCACACCCCTCTCTCCCTCAGGAGGCCCGTGGGCAGACAGGGGAGCC 180

```

```

752 GGGCTTTGCCCTGCTGCTCTCTGCTGATGACCCCTATTAGTATCAGTAGCCACT 811
181 GGGCTTTGCCCTGCTGCTCTCTGCTGATGACCCCTATTAGTATCAGTAGCCACT 240
812 ACTCCCCCTGCTGGCCCTGAGAGCGCTCTGCTGCTATCTCAAGAGAGCCCTGTGCCCA 871
241 ACTCCCCCTGCTGGCCCTGAGAGCGCTCTGCTGCTATCTCAAGAGAGCCCTGTGCCCA 300
872 GCCGTCACCCCTGAGTGGTCTCTTTCAGCTCTTTCCAGCCAGCCAGCCAGCTGCTACGA 931
301 GCCGTCACCCCTGAGTGGTCTCTTTCAGCTCTTTCCAGCCAGCCAGCCAGCTGCTACGA 360
932 CCCCCAGATGTCGCCAAGCACTGTCTCACCAATCCCGCAGCCACCAAGAGAGCCAGG 991
361 CCCCCAGATGTCGCCAAGCACTGTCTCACCAATCCCGCAGCCACCAAGAGAGCCAGG 420
992 GCTGGAGTCAGGCGCACTTTCAGCTGCTCTTCTCGTGCATCGTCTCTCTCTCTCTGCT 1051
421 GCTGGAGTCAGGCGCACTTTCAGCTGCTCTTCTCGTGCATCGTCTCTCTCTCTCTGCT 480
1052 TTTTCTCTCTCCCCCACTTCTCTCTCTGACCCCTCCCTCCGGTGGCTGCTATCA 1111
481 TTTTCTCTCTCCCCCACTTCTCTCTCTGACCCCTCCCTCCGGTGGCTGCTATCA 540
1112 AAGCTCTCAAAACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1171
541 AAGCTCTCAAAACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
1172 TCCTAAGCTATCCAAAGGATGGACCCAGGCTCTGCTGGGAGGTTCCACCCCTTGGATCCAG 1231
601 TCCTAAGCTATCCAAAGGATGGACCCAGGCTCTGCTGGGAGGTTCCACCCCTTGGATCCAG 660
1232 GAAGAACCCCTCCACCCCTGCTGCTGGTGGCCAAAGCTACAGGCTGCTTCTTCTCTCT 1291
661 GAAGAACCCCTCCACCCCTGCTGCTGGTGGCCAAAGCTACAGGCTGCTTCTTCTCTCT 720
1292 CCCCCACCCC 1301
721 CCCCCACCCC 730

RESULT 14
BG769088
LOCUS 602743340F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4873185 5',
mRNA sequence.
ACCESSION BG769088
VERSION BG769088.1 GI:14079741
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 749)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1750 row: 0 column: 10
High quality sequence stop: 745.
Location/Qualifiers
1..749
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source

```

```

/clone="IMAGE:4873185"
/tissue_lib="NIH_MGC_49"
/lab_host="DH10B (phage-resistant)"
/note="organ: skin; Vector: pOTB7; Site:1: XhoI; Site:2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-CDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT      190 a 192 c 220 g 147 t
ORIGIN
Query Match      22.2%; Score 721.8; DB 12; Length 749;
Best Local Similarity 99.5%; Pred. No. 1.3e-114;
Matches 745; Conservative 0; Mismatches 2; Indels 2; Gaps 2;
QY 1339 GACCTGCGAAGTGGAGCATCGAGTAGAGGGAAACAGCAACCGGGAGTCTCGAGCC 1398
Db 1 GACCTGCGAAGTGGAGCATCGAGTAGAGGGAAACAGCAACCGGGAGTCTCGAGCC 60
QY 1399 TGGGCTGCCCTACCTCTACCATTCGCCGACGAGCTTGGCTTGGCTGCCCG 1458
Db 61 TGGGCTGCCCTACCTCTACCATTCGCCGACGAGCTTGGCTTGGCTGCCCG 120
QY 1459 CTTGCTCTTTGGGAACTGAGCTCAGAGGAGGTCTTCAGAGGAAGAAACAAATGAG 1518
Db 121 CTTGCTCTTTGGGAACTGAGCTCAGAGGAGGTCTTCAGAGGAAGAAACAAATGAG 180
QY 1519 GGGTGGCAGGATATAAAGTCACTTCTCTACCTCCCATGAGCATGAACAAATT 1578
Db 181 GGGTGGCAGGATATAAAGTCACTTCTCTACCTCCCATGAGCATGAACAAATT 240
QY 1579 TCTCTCCACCTGGCTCCCAATTTAAAGATGGACCAAGGCTGTGGTACTCCAGGG 1638
Db 241 TCTCTCCACCTGGCTCCCAATTTAAAGATGGACCAAGGCTGTGGTACTCCAGGG 300
QY 1639 CAAGGAGAGCCCTGGGGTCACTGACATGTCTAGGCGCAACCATGCACTCCACAAGGGAG 1698
Db 301 CAAGGAGAGCCCTGGGGTCACTGACATGTCTAGGCGCAACCATGCACTCCACAAGGGAG 360
QY 1699 CATTGGAAATGAAGACTAGTCTCTATGTATFAGTTAAGCAAGGAGAGCTGGCCA 1758
Db 361 CATTGGAAATGAAGACTAGTCTCTATGTATFAGTTAAGCAAGGAGAGCTGGCCA 420
QY 1759 GGGCAGCAGTTTGCACAGCAGAGGGGAATGTAGCAACAGCAGGCGCTCTAGGCCCAT 1818
Db 421 GGGCAGCAGTTTGCACAGCAGAGGGGAATGTAGCAACAGCAGGCGCTCTAGGCCCAT 480
QY 1819 CTTCCATTTCTTAGTAAAGAGAGCATTTCTCTAGACTCCCAAGCGGAGACTGAGCCTA 1878
Db 481 CTTCCATTTCTTAGTAAAGAGAGCATTTCTCTAGACTCCCAAGCGGAGACTGAGCCTA 540
QY 1879 GCCTTCAGCAACCAAGTTCTCTCTGGGACCCAAAGTTTATGGGAAGGCGCAAGACTTC 1938
Db 541 GCCTTCAGCAACCAAGTTCTCTCTGGGACCCAAAGTTTATGGGAAGGCGCAAGACTTC 600
QY 1939 ATGGGAAGAGAGAA -GGAAGGCCCTGGGTAGAAAGCGTTGGT -GCTGTCTCTTTGGCT 1996
Db 601 ATGGGAAGAGAGAAAGGAGGCCCTGGGTAGAAAGCGTTGGTGGTGTCTCTTTGGCT 660
QY 1997 TTAAGCAAAAGGCTCATCTTCCCTCTACCTCTCTATAGGCTTGAGGGTTTGCCAAACA 2056
Db 661 TTAAGCAAAAGGCTCATCTTCCCTCTACCTCTCTATAGGCTTGAGGGTTTGCCAAACA 720
QY 2057 CACTGTGGCTACAGTGGAGGAGAGGA 2085
Db 721 CACTGTGGCTACAGTGGAGGAGAGGA 749

```

RESULT 15
 BM049329
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: DCTB/DTP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LILCM1945 row: m column: 20
 High quality sequence stop: 758.
 Location/Qualifiers
 1..766
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5452987"
 /tissue_lib="NIH_MGC_40"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="organ: prostate; Vector: pOTB7; Site:1: XhoI;
 Site:2: EcoRI; CDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAGCAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-CDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

BASE COUNT 190 a 202 c 221 g 153 t

Query Match 22.0%; Score 716.8; DB 13; Length 766;
 Best Local Similarity 99.2%; Pred. No. 9.3e-114;
 Matches 762; Conservative 0; Mismatches 2; Indels 4; Gaps 4;

QY 1312 CATGTGCCATGGGCTGCTCCCTCCAGTGACCTGCGAAGTGGAGCATCGAGGTAGGAGG 1371
 Db 2 CATGTGCCATGGGCTGCTCCCTCCAGTGACCTGCGAAGTGGAGCATCGAGGTAGGAGG 61
 QY 1372 AAACAGCAACCGGGAGTCTCGAGCTGGGCTGCCCTTACCTTACCATTTCCCGAACC 1431
 Db 62 AAACAGCAACCGGGAGTCTCGAGCTGGGCTGCCCTTACCTTACCATTTCCCGAACC 121
 QY 1432 AGAGCTTTGCCCTTGGCTGCCCTGCTCTTTGGGGAACCTAGACTCAGAGGCAG 1491
 Db 122 AGAGCTTTGCCCTTGGCTGCCCTGCTCTTTGGGGAACCTAGACTCAGAGGCAG 181
 QY 1492 GTGCTTCAGAGAAGAAACAAATAGAGGGTGGCGGGATAGAGGATACCTCCATTCCTC 1551
 Db 182 GTGCTTCAGAGAAGAAACAAATAGAGGGTGGCGGGATAGAGGATACCTCCATTCCTC 241
 QY 1552 TACCTCCCATGAGCATGAACACAATTTCTCTCCAGCTGGCTCCCAAAATTTAAAGATGTG 1611
 Db 242 TACCTCCCATGAGCATGAACACAATTTCTCTCCAGCTGGCTCCCAAAATTTAAAGATGTG 301
 QY 1612 GACCAAGGCTGTGGGTACTCCAGGGGCAAGGAGGCCCTGGGTGAGTCACTGACTCTCAG 1671
 Db 302 GACCAAGGCTGTGGGTACTCCA -GGGCAAGGAGAGGCCCTGGGTGAGTCACTGACTCTCAG 360

```
QY 1672 GCCAACCATGCTCCACAAAGGGAGGAGCATTTGGAAATGAAGGACTAGCTCCTATGTATC 1731
Db |||||||
QY 1732 AGGTTAAGAGCAAGGAGAGCTGGCCAGGAGCAGAGTTTGCACAGCAGAGGGGAATGTA 1791
Db |||||||
QY 1792 GGAACAGCAGGCGCTCCTAGGCGCCCATCTTCCATTTCTTAGGTTAAGAGAGCATTTCCCTC 1851
Db |||||||
QY 1852 AGACTCCCAGGCGGAGGAGCTAGCCCTTACGCAACCAAGGTTCTCTCGGAGCCCAA 1911
Db |||||||
QY 1912 AGTTTATGGAGAGGCGCAAGACTTCATGGGAAGAGAGAGGAGGCGCTGGGTAGAAA 1971
Db |||||||
QY 1972 CGCTTGGTGTCTTCTTTGGCCTTTAAGACAAAGCGCTCATCTTGGCCCTCTAGCTCCT 2031
Db |||||||
QY 2032 GATAGGCTTGAGGTTTGCCAAACCACTGTGGCTACAGGTGAGGGA 2079
Db |||||||
QY 2079 GATAGGCTTGAGG-TTGGCAACCACTGTGGCTACAGGTGAGGGA 2137
Db |||||||
```

Search completed: June 23, 2003, 17:08:10
Job time : 4320 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	156.6	4.8	1340	2	US-08-824-873-2	Sequence 2, Appli
2	156.6	4.8	1340	3	US-09-198-181-2	Sequence 2, Appli
3	156	4.8	925	3	US-08-916-901-4	Sequence 4, Appli
4	156	4.8	925	4	US-09-154-602-4	Sequence 4, Appli
5	136.2	4.2	875	4	US-09-075-454-10	Sequence 10, Appli
6	136.2	4.2	2612	4	US-09-484-970B-142	Sequence 142, Appli
7	120.6	3.7	847	2	US-08-773-423-4	Sequence 4, Appli
8	113	3.5	639	4	US-09-399-913-66	Sequence 66, Appli
9	110	3.4	970	3	US-08-888-077A-28	Sequence 28, Appli
10	91.2	2.8	820	3	US-08-741-411-6	Sequence 6, Appli
11	85.8	2.6	7218	1	US-08-232-463-14	Sequence 14, Appli
12	80.6	2.5	890	3	US-08-741-411-4	Sequence 4, Appli
13	80	2.5	1172	4	US-09-075-454-8	Sequence 8, Appli
14	80	2.5	1255	2	US-08-766-551-6	Sequence 6, Appli
15	80	2.5	1533	4	US-09-075-454-11	Sequence 11, Appli
16	79	2.4	848	3	US-08-741-411-2	Sequence 2, Appli
17	77	2.4	1749	4	US-09-149-476-54	Sequence 54, Appli
18	75.6	2.3	1407	4	US-09-493-914-1	Sequence 1, Appli
19	73.6	2.3	803	4	US-09-075-454-13	Sequence 13, Appli
20	71.6	2.2	591	2	US-09-156-979-1	Sequence 1, Appli
21	71.6	2.2	591	4	US-09-387-341-68	Sequence 68, Appli
22	67.2	2.1	702	3	US-08-842-976-2	Sequence 2, Appli
23	67.2	2.1	702	3	US-09-213-397-2	Sequence 2, Appli
24	67.2	2.1	702	3	US-09-416-489-2	Sequence 2, Appli
25	66.2	2.0	331	4	US-09-325-932A-33	Sequence 33, Appli
26	64.8	2.0	615	1	US-08-247-946A-5	Sequence 5, Appli
27	64.8	2.0	615	5	PCR-US95-06420-5	Sequence 5, Appli

QY 57 TACGATGCTGCTTCCGGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACCTGCTG 116
Db 22 TACGAGCTGCGCTTCAAGGTCATGCTGGTGGGGGACTCGGCTGGGGAAGACCTGCTG 81
QY 117 CTGTGC-CGCTTCACCGCAACAGCTTCCACTCCCTCCGACATCTCCACATCGGTG--TT 173
Db 82 CTGGGTGCGATTCAGGATGGCTTCTTCCGGGGGACCTTCATCTCCACGATAGCAT 141
QY 174 GACTTTAAGATGAAGACCATAGAGTAGACGGCATCAAAGTCCGGATACAGATCTGGGAC 233
Db 142 GACTTCCGGAAACAAAGTTCTGGACGTGGATGGTGTGAAGTGAAGCTGACATGTGGGAC 201
QY 234 ACTGAGGGCAGGAGAGATACAGACCATCAAGACAGTACTATCGGGGGCCCGAGGG 293
Db 202 ACAGCTGCTGAGGAGTCCGCGAGTGTACCCATGCTTACCGGATGCTCATGCT 261
QY 294 ATATTTTGGTCTATGACATAGCAGCGAGCTCTTACCAGCATCATGAAGTGGGTC 353
Db 262 CTGCTGCTGCTACGATGTCAACAAAGGCTCTTTGACAACATCCAGGCTGGGCTG 321
QY 354 AGTGAGCTGGATGACTACGACCAAGCGCTCCAGAGATCTTTATTTGGGATAGGCT 413
Db 322 ACCGAGATCCAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 381
QY 414 GATGAGGAGCAAGACGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 473
Db 382 GACTTCCGCAATGAGGCTGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 441
QY 474 GCGATGGACTTTATGAACAAAGTCCGCTGCGACCAACCTCAACATTAAGAGTCAAT 533
Db 442 GGACTGCGCTTACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 501
QY 534 CGTCTGACAGAGCTGCTGCTGAGGCGCCCATAGGAAGG 570
Db 502 GCCATAGCAAGGAGTGAAGCAGCGCTCCATGAAGG 538

RESULT 2

US-09-198-184-2

; Sequence 2, Application US/09198184

; Patent No. 6010859

; GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Guebler, Karl

; TITLE OF INVENTION: NOVEL RAB PROTEIN

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/198,184

; FILING DATE:

; CLASSIFICATION:

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 08/824,873

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0240 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1340 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: PANCNOT04

; CLONE: 738957

; US-09-198-184-2

Query Match

Best Local Similarity 4.8%; Score 156.6; DB 3; Length 1340;

Matches 305; Conservative 0; Mismatches 209; Indels 3; Gaps 2;

QY 57 TACGATGCTGCTTCCGGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACCTGCTG 116

Db 22 TACGAGCTGCGCTTCAAGGTCATGCTGGTGGGGGACTCGGCTGGGGAAGACCTGCTG 81

QY 117 CTGTGC-CGCTTCACCGCAACAGTTCCTCCCTCCGACATCTCCACATCGGTG--TT 173

Db 82 CTGGGTGCGATTCAGGATGGCTTCTTCCGGGGGACCTTCATCTCCACGATAGCAT 141

QY 174 GACTTTAAGATGAAGACCATAGAGTAGACGGCATCAAAGTCCGGATACAGATCTGGGAC 233

Db 142 GACTTCCGGAAACAAAGTTCTGGACGTGGATGGTGTGAAGTGAAGCTGACATGTGGGAC 201

QY 234 ACTGAGGGCAGGAGAGATACAGACCATCAAGACAGTACTATCGGGGGCCCGAGGG 293

Db 202 ACAGCTGCTGAGGAGTCCGCGAGTGTACCCATGCTTACCGGATGCTCATGCT 261

QY 294 ATATTTTGGTCTATGACATAGCAGCGAGCTCTTACCAGCATCATGAAGTGGGTC 353

Db 262 CTGCTGCTGCTACGATGTCAACAAAGGCTCTTTGACAACATCCAGGCTGGGCTG 321

QY 354 AGTGAGCTGGATGATACGACCAAGCGCTTCCAGAGATCTTTATTTGGGATAGGCT 413

Db 322 ACCGAGATCCAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 381

QY 414 GATGAGGAGCAAGACGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 473

Db 382 GACTTCCGCAATGAGGCTGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 441

QY 474 GCGATGGACTTTATGAACAAAGTCCCTGCGACCAACCTCAACATTAAGAGTCAAT 533

Db 442 GGACTGCGCTTACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 501

QY 534 CGTCTGACAGAGCTGCTGCTGAGGCGCCCATAGGAAGG 570

Db 502 GCCATAGCAAGGAGTGAAGCAGCGCTCCATGAAGG 538

RESULT 3

US-08-916-901-4

; Sequence 4, Application US/08916901

; Patent No. 5892012

; GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Lal, Preeti

; APPLICANT: Corley, Neil C.

; APPLICANT: Shah, Purvi

; TITLE OF INVENTION: RAB PROTEINS

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Dr.

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

```

; SOFTWARE: FastSeq for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/916,901
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
;
; APPLICATION NUMBER:
;
; FILING DATE:
;
; ATTORNEY/AGENT INFORMATION:
;
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0367 US
; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
;
; INFORMATION FOR SEQ ID NO: 4:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 925 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; IMMEDIATE SOURCE:
;
; LIBRARY: LIVRUT04
; CLONE: 2514506
;
; US-08-916-901-4

```

Query Match	4.98;	Score 156;	DB 2:	Length 925;
Best Local Similarity	56.7%;	Pred. No.	9.7e-30;	
Matches	288;	Conservative	0;	Mismatches 220; Indels 0; Gaps 0;

QY	36	GCCCCAGTCATGCGGAAGCAGTACGATGTCTCTTCCGGCTGCTGCTGATCGGGGACTCC	95
Db	57	GCGCGGCCCATGAACCCCGAANTATGACTCTCTTTAAAGCTGTTTTGATTGGCGACTCA	116
QY	96	GGGGTGGGAAGAACCCTCCTCTGCTGTGTCGGCTTCCACGCACAACGAGTTCCACTCCTCGCAC	155
Db	117	GGCGTGGCAAGTCATGCGCTCTCCTCGGCTTGCTGATCACAGTACACAGAGAGCTAC	176
QY	156	ATCTCCACCATCCGGTGTGTGACTTTTAGATGAAGACCATAGAGGTAGACGGCATCAAAGTG	215
Db	177	ATCAGACCACTCCGGGTGGACTTTCAGATCCGAACCATCGAGCTGGATGGCAAAACTATC	236
QY	216	CGGATACAGATCTGGACACTGCAGGCGCAGAGAGATACCAGACCATCACAAAGCAGTAC	275
Db	237	AAACTTCAGATCTGGGACACAGCGGGCCAGGAACGGTTCGGACCACTACTTCCAGCTAC	296
QY	276	TATCGCGGGGCCACAGGGGATATTTTGGTCTATGACATTAGCAGCGAGCGCTCTTACCAG	335
Db	297	TACCGGGGGCTCATGSCATCATCGTGGTCTATGAGCTCACTGACCAGGAATCCTAGCCC	356
QY	336	CACATCATGAAGTGGGTGAGTGAGTGGATGAGTACGCCACAGAACGGCTCCAGAAATC	395
Db	357	NAGCTGAAGCAGTGGCTGCAAGGAGATTGACCGCTATGCCAGCGAAGAACGTCAATAAGCTC	416
QY	396	CTTATTGGGAATAAGCTGTATGAGGAGCAAAAACGGCAGCTGGGAAGAGAGCAGGCGCAG	455
Db	417	CTGGTGGCCNACNAGAGCGHACTCACCAACGAGAAGTGGTGGCAACAACACCCACAGCCAG	476
QY	456	CAGCTCGCGGAAGGAGTATGSGCATGGACTTCTATGAACAAGTGGCTGCACCAACCTCAAC	515
Db	477	GAGTTTGCAGACTCTCTGGGCATCCCTTCTTTGAGACGAGCGCCCAAGATGCCACCAAT	536
QY	516	ATTAAGAGTCATTCAGCGGCTCTGACAG	543
Db	537	GTCCGACGCGGCTTCATGACATGGGTG	564

RESULT 4
US-09-154-602-4
; Sequence 4, Application US/09154602
; Patent No. 6300472
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti

APPLICANT: Corley, Neil C.
 APPLICANT: Shah, Purvi
 TITLE OF INVENTION: RAB PROTEINS
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Dr.
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/154,602
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/916,901
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0367 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 925 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: LIVRTUT04
 CLONE: 2514506
 US-09-154-602-4

	Query Match:	4.8%;	Score 156;	DB 4;	Length 925;	
	Best Local Similarity	56.7%;	Pred. No. 9,7e-30;			
Matches	Conservative	0;	Mismatches	220;	Gaps 0;	
QY	36	GC	CC	CAGTCA	TGCGGAAGCAGTAGTACGATGTCTGTTCCGGCTGCTGCTGATGTCGGGGACTCC	95
Db	57	GC	CGCGCCCATGA	ACCCCAGTAATATGACTACTCTGTTAAAGCTGCTTTTGATTGGGACTCA	116	
QY	96	GGGGTGGGCAAG	ACCCTGCCCTGCTGTGTGGCGGCTTCA	CCGACAACGAGATPCCACTCCTCGCAC	155	
Db	117	GGCGTGGGCAAGT	CATGCCCTGCTCTCGGTTTGCTGATGACACG	TACACAGAGAGCTAC	176	
QY	156	ATCTCCA	CACCATCGGTGTGACTTTAAGATCA	AAGACCATAGAGGTAGACGCGCATCAAGTG	215	
Db	177	ATCAGC	CACCATCGGGTGGACTTCMAGTCCG	AAACCATCGAGTGGTGGCAAACTATC	236	
QY	216	CGGATACAGATCTGGG	ACACTGCGAGGCGAGGAGATAC	CAGACCATATCAAAAGCAGTAC	275	
Db	237	AAACTTCAGATCTGGG	ACACAGGGGGCGAGAACGGTTC	CGGACCATCACTTCCAGCTAC	296	
QY	276	TATCGGGGGCC	CAGGGGATATTTTGGTCTATGACATTA	GACATTAGCAGCGAGCGCTCTTACCAG	335	
Db	297	TACCGGGGGCTCAT	GGCATCTATCGTGGTGTATGACGTCACT	GACCAAGGAATCCTTAGCC	356	
QY	336	CACATCATGAAGTGGGT	CAGTGCAGTGGATGAGTACG	CACACAGAAGCGGTCCAGAAGATC	395	
Db	357	AACTGAAGCAGTGGCT	TGCAGGAGATTGACCGCTATGCC	ACGAGAACGTCATTAAGCTC	416	
QY	396	CTTATTGGGAATAAG	CGCTGATGAGGACCAAAACGCGCAGTGG	GAAGAGACGAAGGCGAC	455	
Db	417	CTGGTGGGCAAC	AAAGAGCGACCTCA	CCACCAAAGGTTGGTGGACRACAC	CCACGACCAAG	476
QY	456	CAGCTGGCGAAGG	ATATGCGATGGACTTCTATGAAACAAGT	GGCTGCTSCACCAACCTCAAC	515	

Db 477 GAGTTTGCAGACTCTCTGGGCATCCCTCTTGGAGACGAGCGCCAGAAGTGCACCAAT 536
QY 516 ATTAAGAGTCAATCACGCGTCTGACAG 543
Db 537 GTCGAGCAGCGGCTCATGACCATGGCTG 564

RESULT 5

US-09-075-454-10
; Sequence 10, Application US/09075454
; Patent No. 6391580
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Tang, Y. Tom
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Patterson, Chandra
; APPLICANT: Batra, Sajeev
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: RAS PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/075,454
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/766,551
; FILING DATE: DECEMBER 12, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Certone, Michael C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0168-1 CIP
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 875 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: UCMCLST01
; CLONE: 1528559
; US-09-075-454-10

Query Match 4.2%; Score 136.2; DB 4; Length 875;
Best Local Similarity 57.5%; Pred. No. 8.9e-25;
Matches 265; Conservative 0; Mismatches 193; Indels 3; Gaps 1;
QY 57 TAGCATGTGCTGTCCGGCTGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACCTGCGCTG 116
Db 72 TAGGACCTCAGCGGCAAGGTGATGCTCTGGGAGACACAGCGGCTGGCAAAACATGTTTC 131
QY 117 CTGTGCGGCTTACCCGACACAGGATTCACCTCT---CGCACATCTCCACCATCGGTGT 173
Db 132 CTGATCCCAATTCAAGACGCGGCTTCTGTCCGGAACTTCATAGCCACCGCTCGGCATA 191

QY 174 GACTTTAAGATGAAGACCATAGAGGTAGACGGGATCAAACTGCGGATACAGATCTGGGAC 233
Db 192 GACTTTCAGGAACAAAGGTGGTGGATGGGATGGAGTGAAGCTGCAGATCTGGGAC 251
QY 234 ACTGAGGCGGAGGAGATACACGACCATCACAAAGCAGTACTATCGCGGGGCCAGGGG 293
Db 252 ACCGCTGGGAGGAAACGGTTCGAGCGTCAACCATGCTTATTACAGAGATGTCAGGCC 311
QY 294 ATATTTTGGTCTATGACATTTAGCAGCGAGCGTCTTTACCAGCACATCATGAAGTGGTC 353
Db 312 TTGCTTCTGCTGATGACATCACCAAAATCTCTTTGACACATCATCAGGCGCTGGCTC 371
QY 354 AGTGACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 413
Db 372 ACTGAGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 431
QY 414 GATGAGGAGCAGAAACGCGAGGTGGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAT 473
Db 432 GATATGAGCAGCGAAGAGTATCCGTTCCGAGACGAGAGACCTTTGGCCAGGAGTAC 491
QY 474 GGCATGGACTTCTATGAACAAGTGGCTCGACCAACCTCAA 514
Db 492 GGTGTTCCCTTCTCTGGAGACGAGCGCAAGACTGGCATGAA 532

RESULT 6

US-09-484-970B-142
; Sequence 142, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 142
; LENGTH: 2612
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. 6426186 412477.1CBI
; US-09-484-970B-142

Query Match 4.2%; Score 136.2; DB 4; Length 2612;
Best Local Similarity 57.5%; Pred. No. 1.5e-24;
Matches 265; Conservative 0; Mismatches 193; Indels 3; Gaps 1;
QY 57 TAGCATGTGCTGTCCGGCTGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACCTGCGCTG 116
Db 91 TACGACCTCAGCGGCAAGGTGATGCTCTGGGAGACACAGCGGCTGGCAAAACATGTTTC 150
QY 117 CTGTGCGGCTTACCGACACAGGTTCCACTCTCT---CGCACATCTCCACCATCGGTGT 173
Db 151 CTGATCCCAATTCAAGACGCGGCTTCTGTCCGGAACCTTCATAGCCACCGTCCGCATA 210
QY 174 GACTTTAAGATGAAGACCATAGAGGTAGACGGGATCAAAAGTGGGATACAGATCTGGGAC 233
Db 211 GACTTCAGGAACAAAGGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 270
QY 234 ACTGAGGCGGAGGAGATACGACCATCACAAAGCAGTACTATCGCGGGGCCAGGGG 293
Db 271 ACCGCTGGGAGGAAACGGTTCGAAAGCTCACCCATGCTTATTACAGAGATGCTCAGGCC 330
QY 294 ATATTTTGGTCTATGACATTTAGCAGCGAGCGCTCTTACCAGCACATCATGAAGTGGTC 353
Db 331 TTGCTTCTGCTGATGACATCACCAAAATCTCTTTTCGACAAACATCATCAGGCGCTGGCTC 390
QY 354 AGTGACGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 413

Db 391 ACTGATTCATGAGTATGCCAGAGGGAGCTGGTGTATCATGCTGTAGCAACAAGGCG 450
QY 414 GATGAGGAGCAGAAAGCGGAGGTGGGAAGAGAGCAAGGCGAGCTGGCGAAGAGTAT 473
Db 451 GATATGAGCAGCGAAGAGTGTATCCGTTCCGAAGACGGAGAGACCTTGGCCAGGAGTAC 510
QY 474 GGCATGGACTTCTATGAACAAGTGGCTGCACCAACTCAA 514
Db 511 GGTGTTCCCTTCTCTGGAGACCAGCGCCAAAGACTGGCATGAA 551

RESULT 7

US-08-773-423-4

; Sequence 4, Application US/08773423
; Patent No. 5869291
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; APPLICANT: Bandman, Olga
; TITLE OF INVENTION: NOVEL RAB PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/773,423
; FILING DATE: Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0183 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 847 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus
; US-08-773-423-4

Query Match 3.7%; Score 120.6; DB 2; Length 847;
Best Local Similarity 51.5%; Pred. No. 7.3e-21;
Matches 301; Conservative 0; Mismatches 280; Indels 3; Gaps 1;
QY 22 TCCCGCCCGCTGGCCCGCCAGTCATGCGGAAGCAGTACGATGCTGTTCCGGCTGCTGC 81
Db 60 TCGCGANCAAGATGGGGGAATGGAACCTGAGGAATATATACCTTGTCTTCAAGGTGCTC 119
QY 82 TGATCGGGGACTCCGGGGTGGGCAAGACTGCTGTGTGCGCTTACCGCAACAGT 141
Db 120 TGATCGCGAATCAGGTGTGGGGAAGCAACTACTCTCTCCGATTACGCGCAATGAGT 179
QY 142 TCCACTCCCTCCGCACATCTCCACCATCGGTTGTTGACITTAAGATGAAGACCATAGAGTAG 201

Db 180 TCAGCCACGACACCGCCACCACCATCGGGGTGAGTTCTCCACCCCGCACTGTGATGTGG 239
QY 202 ACGGCATCAAAAGTGGCGATACAGATCTGGGACACTCTGAGGCGCAGAGAGATACCAAGCA 261
Db 240 GCACCGCTGCTGTCAAGGCTCAGATCTGGGACACAGCTGGCCTGGAGCGGTACCGACCA 299
QY 262 TCACAAGCAGTACTACTGCGGGGCCAGGGGATATTTTTGGTCTATGACATTAGCAGCG 321
Db 300 TCACCTCGGGCTACTACTGCTGTGTCAGTGGGGGCCCTCTCTGGTGTGTTGACCTAACCAAGC 359
QY 322 AGCGCTCTTACCAGCACATCATGAAGTGGGTCAAGTGGGTGAGTGGATGATGACGACCAAG 381
Db 360 ACCAGACCTATGCTGTGGGAGGATGGCTGAAGGAGCTCTATGACCATGCTGAAGCCA 419
QY 382 CGCTCCAGAAAGATCCTTATTGGGAATAAGGCTGATGAGGAGCAGAAACGAGGTGGAA 441
Db 420 CGATCGTCGTATGCTCGTGGTAAACAAGTGNACCTCAGCCAGGCGCGGAAGTGCCCA 479
QY 442 GAGAGCAAGGCGCAGCAGCTGGCGAAGGATGATGGCATGGACTTCTATGAACAAGTGCCT 501
Db 480 CTGAGGAGGCCGAATGTCGCTGAAACAATGAGTGTCTCTCTGGAGACCTCAGCCC 539
QY 502 GCACCAACCTCAACATTA--AGAGTCATTTCAGCGCTCTGACAGACTGCTGTGAGG 558
Db 540 TGGACTCTACCAATGTTGAGCTAGCCCTTGAGACTGTCTTGAAGAAATCTTTGCGAAGG 599
QY 559 CCATAGGAAGAGCTGGAAGGCTCGGAGTGGTGGCCAGCAAT 602
Db 600 TGTCGAAGCAGACAGACAGCATCGGACCAATGCCATCACT 643

RESULT 8

US-09-399-913-66
; Sequence 66, Application US/09399913
; Patent No. 6361971
; GENERAL INFORMATION:
; APPLICANT: Rhodes, Kenneth
; APPLICANT: Betty, Maria
; APPLICANT: Ling, Hual-Ping
; APPLICANT: An, Wenqian
; TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR
; FILE REFERENCE: MNI-070CP2
; CURRENT APPLICATION NUMBER: US/09/399,913
; CURRENT FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER: USSN 60/110,277
; EARLIER FILING DATE: 1998-11-30
; EARLIER APPLICATION NUMBER: USSN 60/110,033
; EARLIER FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: USSN 60/109,333
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: USSN 09/298,731
; EARLIER FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: USSN 09/350,614
; EARLIER FILING DATE: 1999-07-03
; EARLIER APPLICATION NUMBER: USSN 09/350,874
; EARLIER FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 66
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Rattus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(636)
US-09-399-913-66

Query Match 3.5%; Score 113; DB 4; Length 639;
Best Local Similarity 52.2%; Pred. No. 5.2e-19;
Matches 251; Conservative 0; Mismatches 230; Indels 0; Gaps 0;
QY 51 AAGCAGTACGATGCTGTTCCGGCTGCTGCTGATCGGGGACTCCGGGGTGGCAAGACC 110

1 ATGGGTACGCTATCTTCAAGTACATCATCTCGCGACACAGGTGTTGGTAAATCG 60
111 TGCCTGCTGTCGCGCTTCCACGACAAAGAGTTCCACTCCTCGGACATCTCCACCATCGGT 170
61 TGTCTATTGCTACAGTTTACAGACAAGAGTTTTCAGCGGTGTCATGACCTCACAATTGGT 120
171 GTTGACTTTAAGATGAAGACCATAGAGGTAGACGCGCATCAAGTGGCGATACAGATCTGG 230
121 GTAGAGTTTGGTGTGCGAATGATACCAATGATGGGAAACAGATAAACTCCAGATCTGG 180
231 GACACTGCGAGGCGAGAGATACAGACCATCAACAAAGCAGTACTATFCGCGGCGCCAG 290
181 GATACAGCAGGCGCAGGAGTCTTTCGTTCTATCAACAAGGTATATACAGAGGTGCGCG 240
291 GGGATATTTTGGTCTATGACATAGACGCGCGCTTTACCAGCAGCATCATGAGTGG 350
241 GGGGCTTTACTAGTGTATGATATTAACAGAGAGACAGCTTCAACCACTTGACAACTGG 300
351 GTCAGTGACGTGGATGAGTACGACACAGAGGCGTCCAGAGATCCTTTATTGGGAATAG 410
301 TTAGAAGACGCGCTCAGCATTCCAATTCACATGTTTCATCATGCTTATTGGAAATAA 360
411 GCTGATGAGGACGACAAACGCGAGGTGGGAGAGAGCAAGGCGACGCTGCGGAGGAG 470
361 AGTGACTTAGAATCTAGGAGAGAGTGAAGAGGAGAGGTGAAGCTTTTTCGACGAGAG 420
471 TATGCGATGGACTTCTATGAACAAGTGCCTGCACCAACCTCAACATTAAAGAGTCATTC 530
421 CATGAGCTATCTTATGGAACACTTCTGCCAAGACTGCTTCTATGATAGAGGAGCAATT 480
531 A 531
481 A 481

RESULT 9

US-08-888-077A-28
; Sequence 28, Application US/08888077A
; Patent No. 6020143
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE AND USES THEREFOR.
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLZ & MENTLIK
; STREET: 600 SOUTH AVENUE WEST
; CITY: WESTFIELD
; STATE: NJ
; COUNTRY: USA
; ZIP: 07090-1497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/888,077A
; FILING DATE: 03-JUL-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,541
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: PALISI, THOMAS M
; REGISTRATION NUMBER: 36,629
; REFERENCE/DOCKET NUMBER: SCHERING 3.0-017 CIP CIP IV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 654-5000
; TELEFAX: (908) 654-7866

INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 970 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..970
; OTHER INFORMATION: /note= "Y2H9"
US-08-888-077A-28

Query Match 3.4%; Score 110; DB 3; Length 970;
Best Local Similarity 52.9%; Pred. No. 3.6e-18;
Matches 236; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

QY 55 AGTACGATGCTCTTCCGGCTGCTGATCGGGGACTCCGGGTGGGCAAGACCTGCC 114
DB 73 AGTACGACTACCTCTTTAAAGTTGCTCTTATTTGGAGATTCTGGTGTGGAAAGAGTAATC 132
QY 115 TGCCTGCGCGCTTCCACGACAAACGAGTTCCACTCCTCGCACATCTCCACCATCGGTGTG 174
DB 133 TCTGTCTCGATTTACTCGAAATGAGTTTAATCTGGAAGCAAGAGCACCATTTGGAGTAG 192
QY 175 ACTTTAGATGAAGACCATAGAGGTAGAGGCGATCAAGTGCGGATACAGATCTGGGACA 234
DB 193 AGTTTGAACAAGAGCATCCAGGTTGATGGAACAAACATAAAGGCGACAGATATGGGACA 252
QY 235 CTGAGGCGAGGAGATACAGACCATCAACAGAGTACTACTCGGCGGCGCCAGGGGA 294
DB 253 CAGAGGCGAGGAGATATCGAGCTATACATCAGCATATATTCGTGGAGCTAGGTG 312
QY 295 TATTTTGGTCTATGACATTTAGCAGCGGCTCTTACCAGCATCATGAAAGTGGTCA 354
DB 313 CCTTATTGGTTATGACATTTGCTAAACATCTCACATATGAAATCTAGAGCGATSGCTGA 372
QY 355 GTGACGTGGATGAGTACGACACAGAGCGCTCCAGAGATCTTATTGGGAATAGGCTG 414
DB 373 AAGAACTGAGAGATCATGCTGATAGTAACATTTATCTGCTTGTGGGCATTAAGATG 432
QY 415 ATGAGGAGCAGAAACGCGAGGTGGGAGAGAGCAAGGCGACAGCTGGCGAAGAGATG 474
DB 433 ATCTACGTCTCTCAGGCGAGTTCTCTACAGATGAAGCAAGAGCTTTTCAGAAAAAGATG 492
QY 475 GCATGGACTTCTATGAACAAGTGGCC 500
DB 493 GTTGTGCTATTCATTGAAACTTCGGCC 518

RESULT 10

US-08-741-411-6
; Sequence 6, Application US/08741411
; Patent No. 6124116
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: NOVEL RAB PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/741,411
; FILING DATE: Herewith

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0139 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 820 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
IMMEDIATE SOURCE:
LIBRARY:
CLONE: Consensus
US-08-741-411-6

Query Match 2.8%; Score 91.2; DB 3; Length 820;
Best Local Similarity 52.3%; Pred. No. 1.8e-13;
Matches 226; Conservative 0; Mismatches 203; Indels 3; Gaps 1;
QY 167 CGGTGTTGACATTTAAGATGAAGACCATAGAGTAGAGCGCATCAAGTGGGATACAGAT 226
DB 218 CGTGGTGGCTTCTTCAAGAGAGGTGGATGGTGGCGCCACCTCTCTGAAGCTTGAGAT 277
QY 227 CTGGGACACTGCGGAGGAGAGATACAGACCATCACAAAGCAGTACTATCGCGGGC 286
DB 278 CTGGGACACAGCTGGCCAGGAGAGTACCACAGCGTCTGCCACCTCTACTTCAGGGTGC 337
QY 287 CCAGGGATATTTTGGTCTATGACATTAAGCAGCGAGGCTCTTACAGCAGCATCAAGAA 346
DB 338 CAACGCTGCGGCTTCTGGTGTACGACATCACAGGAAGGATTCCTCTCAAGGCTCAGCA 397
QY 347 GTGGTGTAGTACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 403
DB 398 GTGGTGTAGTACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 457
QY 404 GAATAGGCTGTAGGAGGAGAGAAACGCGAGGTGGGAGAGAGAGAGAGAGAGAGAGAGAG 463
DB 458 CAACAAGACGACCTCAGCCAGGAGCGGAGGTGACCTTCCAGGAAGGAAGGAGTTGC 517
QY 464 GAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 523
DB 518 CGACAGCCAGAGTGTGTGTTGATGGAAGTTCGCGCAAACTGAACCAACAGGTTTCGGA 577
QY 524 GTCATTCACGCTGTACAGAGCTGGTGTCTCAGCGCCCATAGGAAGGAGCTGGAAGGCT 583
DB 578 GGTGTTCAATACAGTGGCCAGAGAGTACTCCAGAGAGAGAGAGAGAGAGAGAGAGAG 637
QY 584 CCGGATGCTGC 595
DB 638 ACGGGGGATGC 649

RESULT 11
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 INMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-F15
US-08-232-463-14

Query Match 2.6%; Score 85.8; DB 1; Length 7218;
Best Local Similarity 2.3%; Pred. No. 1.1e-11;
Matches 9; Conservative 251; Mismatches 123; Indels 0; Gaps 0;
QY 804 TAGCCACTACTCCCTGCTGCTGCCCTGAGAGCGCTCTGCTGCTCATCTCAAGAGCCGC 863
DB 1060 TTGGGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1119
QY 864 TGTCCCGAGCCGTCACCTGGAGTGTCTTCTTCAGCTGTTTCCCGCAGCCACAGGCC 923
DB 1120 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1179
QY 924 TGCTAGACCCCGAGGATGTCGCGAAGCACTGTCTACCATCTCCGCGCACCCAGACAA 983
DB 1180 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1239
QY 984 CAGCAGCGCTGGATCCAGGCCACTTTCAGCTGCTCTTCTCGTGCATGCTGCTCT 1043
DB 1240 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1299
QY 1044 TCTGCTTTTCTCTCTCTCCCGCTCTCTTCTGACCCCTCCCTCCGGTGGT 1103
DB 1300 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1359
QY 1104 TCGTATCAAGCTCTCAACCCCTCCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1163
DB 1360 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1419
QY 1164 TCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1186
DB 1420 YYYYYYYYYYYYYGTACCAAA 1442

RESULT 12
US-08-741-411-4
Sequence 4, Application US/08741411

Patent No. 6124116
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Au-Young, Janice
TITLE OF INVENTION: NOVEL RAB PROTEINS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/741,411
FILING DATE: Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0139 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 890 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
LIBRARY:
CLONE: Consensus
US-08-741-411-4

Query Match 2.5%; Score 80.6; DB 3; Length 890;
Best Local Similarity 54.0%; Pred. No. 8.4e-11;
Matches 183; Conservative 2; Mismatches 151; Indels 3; Gaps 1;
QY 30 CGCGTGGCCCGGCTGATGGGAGGAGTACGATGCTGTTCCGGCTGCTGCTGATCGGG 89
DB 36 CCGCGCGCCGAGGATGCGAGCCCGCGCACAGGAGCACCTGTACAAAGTTGCTGATGTC 95
QY 90 GACTCCGGGTGGGCAAGACCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 149
DB 96 GACTCCGGGTGGGCAAGACCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 155
QY 150 TCGACATCTCCACCATCGGTGTTGACTTAAAGATGAAGACCATAGAGGTAGACGCATC 209
DB 156 CACTACGGGGCCACCATCGGGTGGACTTCGCCCTCAAGTCTCAACTGGGACAGCAGG 215
QY 210 A---AAGTCGGGTATACAGTCTGGGACACTGGGACGAGGAGATACAGACCATCACA 266
DB 216 ACTCTGGTGGCTGCGAGCTGTGGGACATCGCGGGGCGAGGAGATTTGGCAACATGACC 275
QY 267 AAGCAGTACTATCGCGGGGCCAGGGGATATTTTGTCTATGACATTTAGCAGCGCGC 326
DB 276 CGAGTATCTACAAAGGAGCTGTTGCTGCTTTGTAGTCTTTGATATATCAAGAGTTCC 335
QY 327 TCTTACAGACATCATGAGTGGGTGCTGAGTGGAT 365
DB 336 ACATTTGAGGCAGTCTTAAATGGAAGTGTGATCTGGAT 374

RESULT 13
US-09-075-454-8
Sequence 8, Application US/09075454
Patent No. 6391580
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Tang, Y. Tom
APPLICANT: Lal, Preeti
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
APPLICANT: Patterson, Chandra
APPLICANT: Batra, Sajeev
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: RAS PROTEINS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/075,454
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/766,551
FILING DATE: DECEMBER 12, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Carrone, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0168-1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1172 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: KIDNOT05
CLONE: 627565
US-09-075-454-8

Query Match 2.5%; Score 80; DB 4; Length 1172;
Best Local Similarity 54.4%; Pred. No. 1.3e-10;
Matches 161; Conservative 0; Mismatches 135; Indels 0; Gaps 0;
QY 65 GCTGTCCGGCTGCTGCTGATCGGGGACTCCGGGTGGGCAAGACCTGCTGCTGCGCG 124
DB 178 GATCTCCAAGTCAATTGTGTTGGGGACCTGTCGGTGGGGAAGACTTGCTCATTAATAG 237
QY 125 CTTACCGGACAGGATTCCTCCTCGGCACATCTCCACCATCGGTGTGACTTTAAGAT 184
DB 238 GTTCTGCAAGACACCTTTGATGAAGATTTACAAGGCCACCATTTGGAGTGGACTTCGAGAT 297
QY 185 GAAGACCATAGAGGTAGCGGCATCAAGTGGCGGATACAGATCTGGGACATTCAGGGCA 244
DB 298 GGAAGATTGAGGTGCTGGGCATTCCTTTCAGTTTGGAGTTGGGATACCGCTGGGCA 357
QY 245 GGAGAGATACAGACCATCACAAGACGATCTATCGCGGGGCCAGGGGATATTTTGT 304
DB 358 GGAGAGGTTCAATGATTCATCACTTACTATAGAGGAGCTCAAGCCATCATCATTTGT 417

Db	658	GGACGATTGAGGTGCTGGGCATCCCTTCAGTTGAGCTTGGGATACCGCTGGCA	717
Qy	245	GGAGAGATACCGACCATCAAAAGCAGTACTATCGCGGGCCCGAGGGGATATTTTGGT	304
Db	718	GGAGAGTTCAATGCAATGTCATCAACCTACTATAGAGGCTCAAGCCATCATTTGT	777
Qy	305	CTATGACATTTAGCAGCGAGCGCTCTTACCGACATCATGAAGTGGGTGACGTACG	360
Db	778	CTTCAACCTGAATGATGTGGCATCTCTGGAACATACCAAGCAGTGGCTGGCCGATG	833

Search completed: June 23, 2003, 13:39:29
Job time : 144 secs

Result No.	Query	Score	Length		DB	ID	Description
			Match				
1	100.0	3257	10	US-09-817-198A-1	Sequence 1, Appl		
2	82.8	28770	10	US-09-817-198A-3	Sequence 3, Appl		
3	35.6	2021	9	US-09-764-868-88	Sequence 88, Appl		
4	18.4	601	10	US-09-817-198A-28	Sequence 28, Appl		
5	18.4	601	10	US-09-817-198A-28	Sequence 28, Appl		
6	16.6	601	10	US-09-817-198A-29	Sequence 29, Appl		
7	16.4	566	10	US-09-817-198A-30	Sequence 30, Appl		
8	15.9	601	10	US-09-764-868-507	Sequence 507, Appl		
9	14.7	481	10	US-09-817-198A-31	Sequence 31, Appl		
10	14.7	481	12	US-09-920-300A-303	Sequence 303, Appl		
11	13.5	463	10	US-10-033-528-303	Sequence 303, Appl		
12	13.0	458	9	US-09-964-824B-26	Sequence 26, Appl		
13	9.4	310	9	US-09-918-995-20848	Sequence 20848, A		
14	9.4	310	9	US-10-046-935-1930	Sequence 1930, Ap		
15	9.4	310	9	US-09-878-178-1930	Sequence 1930, Ap		
16	9.4	310	9	US-10-146-502-1930	Sequence 1930, Ap		
17	6.7	624	10	US-09-794-257-9	Sequence 9, Appl		
18	6.7	1161	10	US-09-794-257-7	Sequence 7, Appl		
19	6.7	2497	10	US-09-834-975-879	Sequence 879, Appl		
19	6.7	2497	10	US-09-834-975-885	Sequence 885, Appl		

Qy	1	TGCCGCTGCCGCCGCCAGTTC	CGGCCCGCGCTGGCCCCAGT	CGTATGCGAAGCAGTACG	60
Db	1	TGCCGCTGCCGCCGCCAGTTC	CGGCCCGCGCTGGCCCCAGT	CGTATGCGAAGCAGTACG	60
Qy	61	ATGTGCTGTTCCGGCTGCTG	TATCGGGGACTCGGGGTGGG	CAAGACCTGCCCTGCTGT	120
Db	61	ATGTGCTGTTCCGGCTGCTG	TATCGGGGACTCGGGGTGGG	CAAGACCTGCCCTGCTGT	120
Qy	121	GCGCCTTCACCGACAACGAG	TTCACCTCCTCGCACATCT	CACACATCGGTGTGACTTTA	180
Db	121	GCGCCTTCACCGACAACGAG	TTCACCTCCTCGCACATCT	CACACATCGGTGTGACTTTA	180
Qy	181	AGATGAAGACCATAGAGT	AGCAGGCGATCAAAAGTCG	GCGATACAGATCTGGGAC	240
Db	181	AGATGAAGACCATAGAGT	AGCAGGCGATCAAAAGTCG	GCGATACAGATCTGGGAC	240
Qy	241	GGCAGGAGAGATACACAG	CAACCATCACAAAGCAGT	ACTATCGGGGGCCAGGGAT	300

Db 241 GGCAGGAGATACAGACCATACAAAAGCAGTACTATCGCGGGGCCAGGGGATATTTT 300
QY 301 TGGTCTATACATATTAGCAGCAGGAGCGCTTTACAGACATCATATGATGGGTGCTAGTACG 360
Db 301 TGGTCTATGACATATTAGCAGCAGGAGCGCTTTACAGACATCATATGATGGGTGCTAGTACG 360
QY 361 TGGATGAGTACACACAGCAGGAGCGCTTCCAGAACATCCTTATTGGGAATAGGCTGATGAGG 420
Db 361 TGGATGAGTACACACAGCAGGAGCGCTTCCAGAACATCCTTATTGGGAATAGGCTGATGAGG 420
QY 421 AGCAGAAACGGCAGGTGGGAAGAGAGCAAGGAGCAGCAGCTGCGAAGGAGTATGCGATGG 480
Db 421 AGCAGAAACGGCAGGTGGGAAGAGAGCAAGGAGCAGCAGCTGCGAAGGAGTATGCGATGG 480
QY 481 ACTTCTATGAACACAGTGCCTGCACCAACCTCAACATTAAGAGTCAATTCACGCGCTGGA 540
Db 481 ACTTCTATGAACACAGTGCCTGCACCAACCTCAACATTAAGAGTCAATTCACGCGCTGGA 540
QY 541 CAGAGCTGTGCTGCAGGCGCCATAGGAAGGAGCTGGAAGGCGCTCCGGATGCGTGCACGA 600
Db 541 CAGAGCTGTGCTGCAGGCGCCATAGGAAGGAGCTGGAAGGCGCTCCGGATGCGTGCACGA 600
QY 601 ATGAGTTGGCACTGCGCAGAGCTGGAGGAGGAGGGCAAAACCGAGGGGCCAGCGAACT 660
Db 601 ATGAGTTGGCACTGCGCAGAGCTGGAGGAGGAGGGCAAAACCGAGGGGCCAGCGAACT 660
QY 661 CTTCCAAACCTGCTGGTGTGCTGAGTCTCTGAGTCTGCTGAGTCTGCTGCTGCTCTCT 720
Db 661 CTTCCAAACCTGCTGGTGTGCTGAGTCTCTGAGTCTGCTGAGTCTGCTGCTGCTCTCT 720
QY 721 CTCAGGAGCGCGTGGGCGAGACAGGGGAGCGGGGCTTTGCCCTGCTGCTGCTCTCTCT 780
Db 721 CTCAGGAGCGCGTGGGCGAGACAGGGGAGCGGGGCTTTGCCCTGCTGCTGCTCTCTCT 780
QY 781 GTGATGACCTATTGAGTATCAGTATAGCCACTACTCCCTGCTGCTGCTGCTGCTGCTCT 840
Db 781 GTGATGACCTATTGAGTATCAGTATAGCCACTACTCCCTGCTGCTGCTGCTGCTGCTCT 840
QY 841 CTGCTGTCTATCTAAGCAGCGCTGCTCCGACCGCTGCTCCAGCGCTGCTGCTGCTGCT 900
Db 841 CTGCTGTCTATCTAAGCAGCGCTGCTCCGACCGCTGCTCCAGCGCTGCTGCTGCTGCT 900
QY 901 GCTGTGTTCCCGAGCCAGAGCGCTGCTACGACCGCCCGAGAGTGTGCCGAAGCAGTCTCT 960
Db 901 GCTGTGTTCCCGAGCCAGAGCGCTGCTACGACCGCCCGAGAGTGTGCCGAAGCAGTCTCT 960
QY 961 ACCATCCCGACCCACAGACACAGCAGCGCTGGAGTCCAGGCGCACTTTCAGTGTCTC 1020
Db 961 ACCATCCCGACCCACAGACACAGCAGCGCTGGAGTCCAGGCGCACTTTCAGTGTCTC 1020
QY 1021 CTTTCTCGGTGATCGTCTCTCTCTGCTGCTTTTCTCTCTCTCCCGCACTCTCTTTCTC 1080
Db 1021 CTTTCTCGGTGATCGTCTCTCTCTGCTGCTTTTCTCTCTCTCCCGCACTCTCTTTCTC 1080
QY 1081 TGACCCCTCCCTCGGTGCTTTGGTATCAAGCTCTCAAAACCCCGTCCCGCGTGTGT 1140
Db 1081 TGACCCCTCCCTCGGTGCTTTGGTATCAAGCTCTCAAAACCCCGTCCCGCGTGTGT 1140
QY 1141 CTTGCTGTGCGAGTCTGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200
Db 1141 CTTGCTGTGCGAGTCTGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200
QY 1201 GCTGTGGGAGGTTCACCTTGGATCCAGGAAGACCTCCACCCCTGCTGCTGCTGCTGCTG 1260
Db 1201 GCTGTGGGAGGTTCACCTTGGATCCAGGAAGACCTCCACCCCTGCTGCTGCTGCTGCTG 1260
QY 1261 GGCCAAAGGCTACAGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1320
Db 1261 GGCCAAAGGCTACAGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1320
QY 1321 TGGGCTGCTCTCCGAGTACCTGCGAAGTGGAGATCGAGGTAGGAGGAAACAGCAA 1380
Db 1321 TGGGCTGCTCTCCGAGTACCTGCGAAGTGGAGATCGAGGTAGGAGGAAACAGCAA 1380

QY 1381 CGGGGAGTCTCTGAGCCTTGGGCTGCCCTACTCTACCCATTCCCGACCAGAGCTTTG 1440
Db 1381 CGGGGAGTCTCTGAGCCTTGGGCTGCCCTACTCTACCCATTCCCGACCAGAGCTTTG 1440
QY 1441 CCCTTGTCTGGCTGGCGCTGCCCTCTTTTGGGGAACCTGAGCTCAGAGGAGGTGCTTCA 1500
Db 1441 CCCTTGTCTGGCTGGCGCTGCCCTCTTTTGGGGAACCTGAGCTCAGAGGAGGTGCTTCA 1500
QY 1501 AGAAGGAACAAATAGAGGGTGGCAGGATATAAAGTCAACCTCCATTCTCTACCTCCCA 1560
Db 1501 AGAAGGAACAAATAGAGGGTGGCAGGATATAAAGTCAACCTCCATTCTCTACCTCCCA 1560
QY 1561 TGCAGCATGAACACAAATTTCTCCACCTGGCTCCCAATTTAAAGATGTGACCAAGGC 1620
Db 1561 TGCAGCATGAACACAAATTTCTCCACCTGGCTCCCAATTTAAAGATGTGACCAAGGC 1620
QY 1621 CTGTGGGTACTCCAGGGGCAAGGAGCGCTGGGTCAGTGAACCTGTGAGGCGCAACCAT 1680
Db 1621 CTGTGGGTACTCCAGGGGCAAGGAGCGCTGGGTCAGTGAACCTGTGAGGCGCAACCAT 1680
QY 1681 GCACCTCCAAAGGGGAGCATTTGGAATGAAGGACTAGCTCTCTATGATCAGTTAAGA 1740
Db 1681 GCACCTCCAAAGGGGAGCATTTGGAATGAAGGACTAGCTCTCTATGATCAGTTAAGA 1740
QY 1741 GCAAGGGAGCTGCCAGGGGACAGCAGTTTCCACAGCAGAGGGGAATGTAGCAACAGCA 1800
Db 1741 GCAAGGGAGCTGCCAGGGGACAGCAGTTTCCACAGCAGAGGGGAATGTAGCAACAGCA 1800
QY 1801 GGGCTCTCTAGGCGCCATCTTCCATTTCTTAGTGAAGAGACATTTCTCAGACTCCCA 1860
Db 1801 GGGCTCTCTAGGCGCCATCTTCCATTTCTTAGTGAAGAGACATTTCTCAGACTCCCA 1860
QY 1861 GSGGAGGACTGAGCTTAGCCTTACGCAACCAAGGTTCTCTGGGACCCCAAGTTTATGG 1920
Db 1861 GSGGAGGACTGAGCTTAGCCTTACGCAACCAAGGTTCTCTGGGACCCCAAGTTTATGG 1920
QY 1921 GAGAAGGCAAGACTTCAATGGGAAGAGAGAAAGGCCCTTGGGTAGAAACGCTTGGTG 1980
Db 1921 GAGAAGGCAAGACTTCAATGGGAAGAGAGAAAGGCCCTTGGGTAGAAACGCTTGGTG 1980
QY 1981 CTGTCTCTCTTGGCTTTAAGCAAAAGCGCTCATCTTGCCTCTACCTCTCTGATAGGCTT 2040
Db 1981 CTGTCTCTCTTGGCTTTAAGCAAAAGCGCTCATCTTGCCTCTACCTCTCTGATAGGCTT 2040
QY 2041 GAGGTTTGCACACACTGTGGCTACAGGTGGAGGGAAGAGGACTCTCTCTCCACAG 2100
Db 2041 GAGGTTTGCACACACTGTGGCTACAGGTGGAGGGAAGAGGACTCTCTCTCCACAG 2100
QY 2101 TGCTATGTTAGGAAGTTTCTTTAAGGAGTATGGCCCAAGAGTGTGTTAGGAGGCC 2160
Db 2101 TGCTATGTTAGGAAGTTTCTTTAAGGAGTATGGCCCAAGAGTGTGTTAGGAGGCC 2160
QY 2161 TTTTAAAGCGGAACAAAGTAAATTTACAGTCTCTCTGAGGTTCTCTCCACCGTCCCAAGG 2220
Db 2161 TTTTAAAGCGGAACAAAGTAAATTTACAGTCTCTCTGAGGTTCTCTCCACCGTCCCAAGG 2220
QY 2221 TGGGCGAGGCTTAGGAAGAGGTCATTTAAAGCCACACATTTAGCTGCTGCTGCTG 2280
Db 2221 TGGGCGAGGCTTAGGAAGAGGTCATTTAAAGCCACACATTTAGCTGCTGCTGCTGCTG 2280
QY 2281 CAGCCAAACAAAGAACTGGGTGTGATTTTCACTAATGAACCAAAATTCAGAGGCA 2340
Db 2281 CAGCCAAACAAAGAACTGGGTGTGATTTTCACTAATGAACCAAAATTCAGAGGCA 2340
QY 2341 CTCATATGTGAAGATAGAACCTCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2400
Db 2341 CTCATATGTGAAGATAGAACCTCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2400
QY 2401 ACCATCAAAACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2460
Db 2401 ACCATCAAAACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2460

; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 507
; LENGTH: 566
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (484)
; OTHER INFORMATION: n equals a.t.g, or c
; NAME/KEY: SITE
; LOCATION: (493)
; OTHER INFORMATION: n equals a.t.g, or c
; NAME/KEY: SITE
; LOCATION: (538)
; OTHER INFORMATION: n equals a.t.g, or c
; NAME/KEY: SITE
; LOCATION: (563)
; OTHER INFORMATION: n equals a.t.g, or c
US-09-764-868-507

Query Match 16.4%; Score 534.2; DB 9; Length 566;
Best Local Similarity 98.9%; Pred. No. 2.1e-140;
Matches 533; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 GCCCGCTGCCCGCCGAGTTCCCGCGCGCTGCCCGCCGAGTCATGGCGAAGCAGTACGA 61
DB GCCCGCTGCCCGCCGAGTTCCCGCGCGCTGCCCGCCGAGTCATGGCGAAGCAGTACGA 87
QY 62 TGTGCTGTTCCCGCTGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACCTGCTGCTGTG 121
DB TGTGCTGTTCCCGCTGCTGCTGATCGGGGACTCCGGGGTGGGCAAGACCTGCTGCTGTG 147
QY 122 CCGCTTCACGACACAGTTCACATCTCTCGCATCTCCACCATCGGTGTTGACTTTAA 181
DB CCGCTTCACGACACAGTTCACATCTCTCGCATCTCCACCATCGGTGTTGACTTTAA 207
QY 182 GATGAGACCATAGAGGTAGACGGCATCAAGTGGGATACACATCTGGGACACTGCAGG 241
DB GATGAGACCATAGAGGTAGACGGCATCAAGTGGGATACACATCTGGGACACTGCAGG 267
QY 242 GCAGGAGATACAGACCATCAAGAGTACTATCGCGGGCCGAGGGGATATTTT 301
DB GCAGGAGATACAGACCATCAAGAGTACTATCGCGGGCCGAGGGGATATTTT 327
QY 302 GGTCTATGACATTAGCAGCGAGCGCTCTTACCAGCACATCATGAAGTGGTCACTGACGT 361
DB GGTCTATGACATTAGCAGCGAGCGCTCTTACCAGCACATCATGAAGTGGTCACTGACGT 387
QY 362 GGATGAGTACGACAGAGAGCGCTCCAGAGATCTTATGGGAATAAGCTCATGAGGA 421
DB GGATGAGTACGACAGAGAGCGCTCCAGAGATCTTATGGGAATAAGCTCATGAGGA 447
QY 422 GCAGAAACGCGAGGTGGGAAGAGAGCAAGGCGAGCAGCTGGCGAAGGAGTATGGCATGGA 481
DB GCAGAAACGCGAGGTGGGAAGAGAGCAAGGCGAGCAGCTGGCGAAGGAGTATGGCATGGA 507
QY 482 CTCTATGAACAAGTGCCTGCGACCACTCAACATTAAGATGTCATCAAGCTGCTGA 540
DB CTCTATGAACAAGTGCCTGCGACCACTCAACATTAAGATGTCATCAAGCTGCTGA 566

RESULT 8

US-09-817-198A-31
; Sequence 31, Application US/09817198A
; Patent No. US20020146758A1
; GENERAL INFORMATION:

; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: C1001188
; CURRENT APPLICATION NUMBER: US/09/817,198A

; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-817-198A-31

Query Match 15.9%; Score 516.6; DB 10; Length 601;
Best Local Similarity 99.8%; Pred. No. 2e-135;
Matches 516; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2707 GCAGTGGCTCCCAATCTGTGGACCACTATTTCAGCTTTCCTTCAAGATCAGGAGGTGC 2766
DB 1 GGCAGTGGCTCCCAATCTGTGGACCACTATTTCAGCTTTCCTTCAAGATCAGGAGGTGC 60
QY 2767 CATTCAATGCTTTCTCTCTAGCCCCCTCAGCAAGAGGACTATATTTGTACTGTACC 2826
DB 61 CATTCAATGCTTTCTCTCTAGCCCCCTCAGCAAGAGGACTATATTTGTACTGTACC 120
QY 2827 CTAGGGGTTCTGGAAGGAAACATGGAATCAGGATTTCTATAGACTGATAGGCTTATCC 2886
DB 121 CTAGGGGTTCTGGAAGGAAACATGGAATCAGGATTTCTATAGACTGATAGGCTTATCC 180
QY 2887 ACAAGGGCCATGACTGGGAAAAGGTATGGGAGCAGAGAGAATTTGGGATTTAGGGTGC 2946
DB 181 ACAAGGGCCATGACTGGGAAAAGGTATGGGAGCAGAGAGAATTTGGGATTTAGGGTGC 240
QY 2947 AGCTAGCTCACCTCAACCTTTTGGTGGCTGGGGCATGCTTGAGGCCCCAGACTGTAA 3006
DB 241 AGCTAGCTCACCTCAACCTTTTGGTGGCTGGGGCATGCTTGAGGCCCCAGACTGTAA 300
QY 3007 GCAGGCTCTGCTGGCTGTTTACTCGTCACCACTCTGCACTGCTGCTGTGAGACTCCA 3066
DB 301 GCAGGCTCTGCTGGCTGTTTACTCGTCACCACTCTGCACTGCTGCTGTGAGACTCCA 360
QY 3067 TCAGGCCCCAGGCGACCCCTGCTCTGAGCCTTCCACTATCTCCCTGTGACGGGTGAAC 3126
DB 361 TCAGGCCCCAGGCGACCCCTGCTCTGAGCCTTCCACTATCTCCCTGTGACGGGTGAAC 420
QY 3127 TTCTGTACTGCTCTCGGCTCCATATATGTAATTTGAGCAGGGTTCATCTATTTTAAAC 3186
DB 421 TTCTGTACTGCTCTCGGCTCCATATATGTAATTTGAGCAGGGTTCATCTATTTTAAAC 480
QY 3187 ACAGATGTTTACAAAATAAGATTATTTCAACCACC 3223
DB 481 ACAGATGTTTACAAAATAAGATTATTTCAACCACC 517

RESULT 9

US-09-920-300A-303
; Sequence 303, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:

; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547
; CURRENT APPLICATION NUMBER: US/09/920,300A
; CURRENT FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 303
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 417, 461

; OTHER INFORMATION: n = A,T,C or G
US-09-920-300A-303

Query Match 14.7%; Score 477.4; DB 10; Length 481;
Best Local Similarity 99.4%; Pred. No. 2.1e-124;
Matches 478; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2582 GTTGCTTCCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATCAGAGTTA 2701
DB 1 GTTGCTTCCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATCAGAGTTA 120
QY 2642 CTAGGGAAGTAGGTAGTCCCAATCTGTGGACCAAGTATTTTCCCTGAAGATCAGGCAG 2761
DB 61 CTAGGGAAGTAGGTAGTCCCAATCTGTGGACCAAGTATTTTCCCTGAAGATCAGGCAG 120
QY 2702 TATTTGGGCACTGGCTCCCAATCTGTGGACCAAGTATTTTCCCTGAAGATCAGGCAG 180
DB 121 TATTTGGGCACTGGCTCCCAATCTGTGGACCAAGTATTTTCCCTGAAGATCAGGCAG 2761
QY 2762 GGTGCCATTATCTCTTCTCTAGCCCTCAGGAAAGAGGACTATATTTGTACT 2821
DB 181 GGTGCCATTATCTCTTCTCTAGCCCTCAGGAAAGAGGACTATATTTGTACT 240
QY 2822 GTACCTAGGGGTTCTGGAAGGAAACATGGAATCAGGATTTATAGACTGATAGGCC 2881
DB 241 GTACCTAGGGGTTCTGGAAGGAAACATGGAATCAGGATTTATAGACTGATAGGCC 300
QY 2882 TATCCACAAGGGCCATGACTGGGAAAGGATGAGGAGAGGAAATGGGATTTAG 2941
DB 301 TATCCACAAGGGCCATGACTGGGAAAGGATGAGGAGAGGAAATGGGATTTAG 360
QY 2942 GGTGCAGCTAGCTCACCCTAAACTTTTGGTGGCTGGGCGATGCTTGAGGCCAGACT 3001
DB 361 GGTGCAGCTAGCTCACCCTAAACTTTTGGTGGCTGGGCGATGCTTGAGGCCAGACT 420
QY 3002 GTTAAGCAGGCTCTGCTGGCCTGTTTACTGCTCACCACCTGTCACCTGCTGTCTTGAGA 3061
DB 421 GTTAAGCAGGCTCTGCTGGCCTGTTTACTGCTCACCACCTGTCACCTGCTGTCTTGAGA 480
QY 3062 C 3062
DB 481 C 481

RESULT 10
US-10-033-528-303
; Sequence 303, Application US/10033528
; Patent No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secret, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 303
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 417, 461
; OTHER INFORMATION: n = A,T,C or G
US-10-033-528-303

Query Match 14.7%; Score 477.4; DB 12; Length 481;
Best Local Similarity 99.4%; Pred. No. 2.1e-124;
Matches 478; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2582 GTTGCTTCCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATCAGAGTTA 2641
DB 1 GTTGCTTCCTTGAAGATGGGACTCCTTGGGTATCAAGACCTATGCCACATCAGAGTTA 60
QY 2642 CTAGGGAAGTAGGTAGTCCCAATCTGTGGACCAAGTATTTTCCCTGAAGATCAGGCAG 2701
DB 61 CTAGGGAAGTAGGTAGTCCCAATCTGTGGACCAAGTATTTTCCCTGAAGATCAGGCAG 120
QY 2702 TATTTGGGCACTGGCTCCCAATCTGTGGACCAAGTATTTTCCCTGAAGATCAGGCAG 180
DB 121 TATTTGGGCACTGGCTCCCAATCTGTGGACCAAGTATTTTCCCTGAAGATCAGGCAG 2761
QY 2762 GGTGCCATTATCTCTTCTCTAGCCCTCAGGAAAGAGGACTATATTTGTACT 2821
DB 181 GGTGCCATTATCTCTTCTCTAGCCCTCAGGAAAGAGGACTATATTTGTACT 240
QY 2822 GTACCTAGGGGTTCTGGAAGGAAACATGGAATCAGGATTTATAGACTGATAGGCC 2881
DB 241 GTACCTAGGGGTTCTGGAAGGAAACATGGAATCAGGATTTATAGACTGATAGGCC 300
QY 2882 TATCCACAAGGGCCATGACTGGGAAAGGATGAGGAGAGGAAATGGGATTTAG 2941
DB 301 TATCCACAAGGGCCATGACTGGGAAAGGATGAGGAGAGGAAATGGGATTTAG 360
QY 2942 GGTGCAGCTAGCTCACCCTAAACTTTTGGTGGCTGGGCGATGCTTGAGGCCAGACT 3001
DB 361 GGTGCAGCTAGCTCACCCTAAACTTTTGGTGGCTGGGCGATGCTTGAGGCCAGACT 420
QY 3002 GTTAAGCAGGCTCTGCTGGCCTGTTTACTGCTCACCACCTGTCACCTGCTGTCTTGAGA 3061
DB 421 GTTAAGCAGGCTCTGCTGGCCTGTTTACTGCTCACCACCTGTCACCTGCTGTCTTGAGA 480
QY 3062 C 3062
DB 481 C 481

RESULT 11
US-09-964-824A-26/c
; Sequence 26, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-26

Query Match 13.5%; Score 438.8; DB 10; Length 453;
Best Local Similarity 99.3%; Pred. No. 1.7e-113;
Matches 451; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2790 CCCCTCAGGAAAGAGGACTATATTTTCTACTGTACCTAGGGTCTTGGAGGAAAC 2849
DB 453 CCCCTCAGGAAAGAGGACTATATTTTCTACTGTACCTAGGGTCTTGGAGGAAAC 394
QY 2850 ATGAATCAGGATTTCTATAGACTGATAGGCCCTATCCACAGGCCCATGACTGGGAAAG 2909
|||||

Db 393 ATGGAATCAGGATCTATAGACTGATAGGCCCTATCCACAAGGGCCATGACTGGGAAAAG 334
Qy 2910 GTATGGAGCAGAGGAGAAATGGGATTTAGGTGCAGTACGCTCACCTAACTTTT 2969
Db 333 GTATGGAGCAGAGGAGAAATGGGATTTAGGTGCAG-TAGGCTCACCTAACTTTT 275
Qy 2970 GGTGGCTGGGGATGCTTTGAGGCCAGACTGTTAGCAGGCTCTGCTGGCCTGTTTAC 3029
Db 274 GGTGGCTGGGGATGCTTTGAGGCCAGACTGTTAGCAGGCTCTGCTGGCCTGTTTAC 215
Qy 3030 TCGTCACCACTCTGCACCTGCTGCTTTGAGACTCCATCCAGGCCAGCGCCACCTG 3089
Db 214 TCGTCACCACTCTGCACCTGCTGCTTTGAGACTCCATCCAGGCCAGCGCCACCTG 155
Qy 3090 CTCTGAGCCTCCACTATCTCCCTGTGACGGGTGAACCTCGTGACTGTGTCGGGTCC 3149
Db 154 CTCTGAGCCTCCACTATCTCCCTGTGACGGGTGAACCTCGTGACTGTGTCGGGTCC 95
Qy 3150 ATATATCAATGTGACGAGGTTCATCTATTTTAAACACAGATGTTTACAAATAAGAT 3209
Db 94 ATATATCAATGTGACGAGGTTCATCTATTTTAAACACAGATGTTTACAAATAAGAT 35
Qy 3210 TATTTCACCAACCAAAAAAATAAAAAA 3243
Db 34 TATTTCACCAACCAAAAAAATAAAAAA 1

RESULT 12

US-09-918-995-20848
; Sequence 20848, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20848
; LENGTH: 458
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(458)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-20848

Query Match 13.0%; Score 424.4; DB 9; Length 458;
Best Local Similarity 99.5%; Pred. No. 1.9e-109;
Matches 425; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2540 AGGACAGCTAACTCTAGGATGAAGCTGTGACTAGGCTGGAGTGTCTCTTGAAGATG 2599
Db 32 AGGACAGCTAACTCTAGGATGAAGCTGTGACTAGGCTGGAGTGTCTCTTGAAGATG 91
Qy 2600 GGACTCTTGGGTATCAAGACCTATGCCACATCACACTGGGGCTAGGGAAGTAGGTGATG 2659
Db 92 GGACTCTTGGGTATCAAGACCTATGCCACATCACACTGGGGCTAGGGAAGTAGGTGATG 151
Qy 2660 CCAGCCCTCAAGTGTGCTTTTCAGCCAGGACTTGAGAAGTTATATGGGCAGTGGCTCCA 2719
Db 152 CCAGCCCTCAAGTGTGCTTTTCAGCCAGGACTTGAGAAGTTATATGGGCAGTGGCTCCA 211
Qy 2720 ATCTGTGGCAGGATTTTTCAGCTTTCCCTGAAGATCAGGAGGCTGCCATTCATTTGCTT 2779
Db 212 ATCTGTGGCAGGATTTTTCAGCTTTCCCTGAAGATCAGGAGGCTGCCATTCATTTGCTT 271
Qy 2780 TCTCTCTAGCCCTCAGGAAAAGAGGACTATATTTGTACTGTACCCCTAGGGGTTCTGG 2839

Db 272 TCCTCTCTAGCCCTCAGGAAAAGAGGACTATATTTGTACTGTACCTTAGGGTTCTGG 331
Qy 2840 AAGGAAAACATGAATCAGGATTTCTATAGACTGATAGGCCCTATCCACAAGGGCCATGA 2899
Db 332 AAGGAAAACATGAATCAGGATTTCTATAGACTGATAGGCCCTATCCACAAGGGCCATGA 391
Qy 2900 CTGGAAAAGGATTTGGGAGCAGAGGAGAAATTTAGGTGTCAGCTACCTCACC 2959
Db 392 CTGGAAAAGGATTTGGGAGCAGAGGAGAAATTTAGGTGTCAGCTACCTCACC 451
Qy 2960 CTAAACT 2966
Db 452 CTAACT 458

RESULT 13

US-10-046-935-1930
; Sequence 1930, Application US/10046935
; Patent No. US20020156011A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secretist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1930
; LENGTH: 310
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-046-935-1930

Query Match 9.4%; Score 306.8; DB 9; Length 310;
Best Local Similarity 99.4%; Pred. No. 2.6e-76;
Matches 308; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2824 ACCCTAGGGTTCTGGAAGGAAAACATGGAATCAGGATTTCTATAGACTGATAGGCCCTA 2883
Db 1 ACCCTAGGGTTCTGGAAGGAAAACATGGAATCAGGATTTCTATAGACTGATAGGCCCTA 60
Qy 2884 TCACAAGGGCCATGACTGCTGGAAAAGGATTTGGGAGCAGAGGAGAAATTTAGG 2943
Db 61 TCACAAGGGCCATGACTGCTGGAAAAGGATTTGGGAGCAGAGGAGAAATTTAGG 120
Qy 2944 TGCACTACGCTCACCTAACCTTTTGTGGCTGGGGCATGTCTTCAGGCCACAGACTGT 3003
Db 121 TGCACTACGCTCACCTAACCTTTTGTGGCTGGGGCATGTCTTCAGGCCACAGACTGT 180
Qy 3004 TAAGCAGGCTCTGCTGGCCTGTTTACTCGTCACACACCTCTGCACCTGCTGTTGAGACT 3063
Db 181 TAAGCAGGCTCTGCTGGCCTGTTTACTCGTCACACACCTCTGCACCTGCTGTTGAGACT 240
Qy 3064 CCATCCAGCCCCCAGGACGCCACCTGCTCTGAGCCCTCCACTATCTCCCTGTGACGGGTG 3123
Db 241 CCATCCAGCCCCCAGGACGCCACCTGCTCTGAGCCCTCCACTATCTCCCTGTGACGGGTG 300
Qy 3124 AACTTCGTGT 3133
Db 301 AACTTCGTGT 310

RESULT 14

US-09-878-178-1930
; Sequence 1930, Application US/09878178
; Patent No. US20020177552A1
; GENERAL INFORMATION:

```
; APPLICANT: Jlang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527
; CURRENT APPLICATION NUMBER: US/09/878,178
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 2237
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1930
; LENGTH: 310
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-878-178-1930

Query Match
Best Local Similarity 99.4%; Score 306.8; DB 9; Length 310;
Matches 308; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2824 ACCCTAGGGTCTCTGGAGGAAACATGGAATCAGGATTTCTATAGACTGATGGCCCTA 2883
Db 1 ACCCTAGGGTCTCTGGAGGAAACATGGAATCAGGATTTCTATAGACTGATGGCCCTA 60

QY 2884 TCCACAAGGCCCATGACTGGGAAAGGATATGGGAGCAGAGGAGAAATTTGGGATTTAGGG 2943
Db 61 TCCACAAGGCCCATGACTGGGAAAGGATATGGGAGCAGAGGAGAAATTTGGGATTTAGGG 120

QY 2944 TGCAGCTACGCTCACCCCTAAACATTTTGTGGCCTTGGGCGCATGTCTTCCCTGTGACGGTG 3003
Db 121 TGCAGCTACGCTCACCCCTAAACATTTTGTGGCCTTGGGCGCATGTCTTCCCTGTGACGGTG 180

QY 3004 TAACGAGGCTCTGCTGGCCTGTTTACTCGTCACCACTCTGCACCTGCTGCTGTGAGACT 3063
Db 181 TAACGAGGCTCTGCTGGCCTGTTTACTCGTCACCACTCTGCACCTGCTGCTGTGAGACT 240

QY 3064 CCATCCAGCCCCAGGCACGCCACCTGCTCTGAGCCTCCACTATCTCCCTGTGACGGTG 3123
Db 241 CCATCCAGCCCCAGGCACGCCACCTGCTCTGAGCCTCCACTATCTCCCTGTGACGGTG 300

QY 3124 AACTTCGTGT 3133
Db 301 AACTTCGTGT 310
```

Search completed: June 23, 2003, 19:46:43
Job time : 445 secs

```
RESULT 15
US-10-146-502-1930
; Sequence 1930, Application US/10146502
; Publication No. US20030069180A1
; GENERAL INFORMATION:
; APPLICANT: Jlang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C2
; CURRENT APPLICATION NUMBER: US/10/146,502
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1930
; LENGTH: 310
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-146-502-1930

Query Match
Best Local Similarity 99.4%; Score 306.8; DB 9; Length 310;
Matches 308; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2824 ACCCTAGGGTCTCTGGAGGAAACATGGAATCAGGATTTCTATAGACTGATGGCCCTA 2883
```

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0233397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0233400.
PR 14-SEP-2000; 2000US-0233401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465460/50.
DR N-PSDB; AAS27053.
XX
PT Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders
XX
PS Claim 1; SEQ ID No 701; 880pp; English.
XX

CC The invention relates to novel isolated polypeptides (I), and
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
 CC diagnosing, preventing and treating diseases including immune system
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
 CC transplant rejections and graft versus host disease, infectious diseases
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative
 CC disorders, primary haematopoietic disorders, hyperproliferative
 CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
 CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
 CC Addison's disease), reproductive system disorders, gastrointestinal
 CC disorder (inflammatory disorders), liver disorders (cirrhosis),
 CC as stimulators of B-cell responsiveness to pathogens, activators of
 CC T-cells, to induce higher affinity antibodies, and as a means to induce
 CC tumour proliferation in pathologies e.g. acquired immune deficiency
 CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction
 CC pathway protein, amino acid sequences of the invention.
 XX

Query Match 100.0%; Score 1105; DB 22; Length 401;
 Best Local Similarity 100.0%; Pred. No. 2.1e-108;
 Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAKQYDVLRLLLIGDSGVGKTCCLCRFTDNEFHSHSTIGVDFKMTIEVDGKVRQ 60
 DB 27 MAKQYDVLRLLLIGDSGVGKTCCLCRFTDNEFHSHSTIGVDFKMTIEVDGKVRQ 86
 QY 61 IWDTAGOERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSVDVEYAPGVOKILIG 120
 DB 87 IWDTAGOERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSVDVEYAPGVOKILIG 146
 QY 121 NKADEOKROVGRGQGLAKEYGMDVETSACTNLNLIKESFTRLTVELVQAHKLEGL 180
 DB 147 NKADEOKROVGRGQGLAKEYGMDVETSACTNLNLIKESFTRLTVELVQAHKLEGL 206
 QY 181 RMASNELALAELEEKGKPEGPANSSKTCWC 212
 DB 207 RMASNELALAELEEKGKPEGPANSSKTCWC 238

RESULT 2
 ID AAB41604 standard; Protein; 218 AA.
 XX AC AAB41604;
 XX DT 08-FEB-2001 (first entry)
 XX DE Human ORF1368 polypeptide sequence SEQ ID NO:2736.
 XX KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antihypertensive;
 KW antinaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; erythematoidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematoidism; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 OS Homo sapiens.
 XX OS
 XX WO200058473-A2.
 PN

XX 05-OCT-2000.
 PD 31-MAR-2000; 2000WO-US08621.
 XX 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX (CURA-) CURAGEN CORP.
 XX Shimkets RA, Leach M;
 XX WPI; 2000-602362/57.
 DR N-PSDB; AAC75813.
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX Claim 11; Page 1979-1980; 5507pp; English.
 XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antifungal; antirheumatic;
 CC antihypertensive; antinaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hyperlension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematoidism, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX SQ Sequence 218 AA;
 Query Match 98.8%; Score 1092; DB 21; Length 218;
 Best Local Similarity 97.2%; Pred. No. 2e-107;
 Matches 212; Conservative 0; Mismatches 0; Indels 6; Gaps 1;
 QY 1 MAKQYDVLRLLLIGDSGVGKTCCLCRFTDNEFHSHSTIGVDFKMTIEVDGKVRQ 60
 DB 1 MAKQYDVLRLLLIGDSGVGKTCCLCRFTDNEFHSHSTIGVDFKMTIEVDGKVRQ 60
 QY 61 IWDTAGOERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSVDVEYAPGVOKILIG 120
 DB 61 IWDTAGOERYQTITKQYRRAQGIPLVYDISSERSYQHIMKWSVDVEYAPGVOKILIG 120
 QY 121 NKADEOKROVGRGQGLAKEYGMDVETSACTNLNLIKESFTRLTVELVQAHK 174
 DB 121 NKADEOKROVGRGQGLAKEYGMDVETSACTNLNLIKESFTRLTVELVQAHK 180
 QY 175 KELEGLRMASNELALAELEEKGKPEGPANSSKTCWC 212
 DB 181 KELEGLRMASNELALAELEEKGKPEGPANSSKTCWC 218
 RESULT 3
 ID AAU17555 standard; Protein; 188 AA.
 XX AC AAU17555;

[illegible]

PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2000US-0254097.
 XX 05-JAN-2001; 2001US-0259678.
 FA (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-465460/50.
 XX N-PSDB; AAS27472.
 DR Novel polypeptides useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cancers, immune
 PT disorders and neuronal disorders -
 XX

Claim 1; SEQ ID No 1120; 880pp; English.

The invention relates to novel isolated polypeptides (I), and polynucleotides (II). (I), (II) and the antibody to (I) are useful for diagnosing, preventing and treating diseases including immune system disorders (e.g. congenital and acquired immunodeficiencies, autoimmune disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ transplant rejections and graft versus host disease, infectious diseases (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and other blood-related disorders (sickle cell anaemia), myeloproliferative disorders, primary haematopoietic disorders, hyperproliferative disorders (e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g. glomerulonephritis), cardiovascular disorders (e.g. arrhythmia), respiratory disorders, dermatological disorders, in wound healing, epithelial cell proliferation, endocrine disorders (e.g. Addison's disease), reproductive system disorders, gastrointestinal disorder (inflammatory disorders), liver disorders (cirrhosis), as stimulators of B-cell responsiveness to pathogens, activators of T-cells, to induce higher affinity antibodies, and as a means to induce tumour proliferation in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction pathway protein, amino acid sequences of the invention.

Query Match 75.3%; Score 832; DB 22; Length 188;
 Best Local Similarity 97.6%; Pred. No. 6.2e-80;
 Matches 160; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMTIEVDGKIRIQ 60
 Db 24 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMTIEVDGKIRIQ 83
 Qy 61 IWDTAGQERYQTITKQYIRRAQGIPLVDYDSSERSYOHIMKWSDVDYAPGKILIG 120
 Db 84 IWDTAGQERYQTITKQYIRRAQGIPLVDYDSSERSYOHIMKWSDVDYAPGKILIG 143
 Qy 121 NKADEQKROVREGQGOOLAKKEYGMDFYETSACTNLNLIKESFTR 164
 Db 144 NKADEQKROVREGQGOOLAKKEYGMDFYETSACTNLNLIKESFTR 187

RESULT 4
 ABB70670
 ID ABB70670 standard; Protein; 204 AA.
 XX

AC ABB70670;
 XX 26-MAR-2002 (first entry)
 DT Drosophila melanogaster polypeptide SEQ ID NO 38802.
 XX Drosophila melanogaster polypeptide; cell signalling; insecticide;
 DE Drosophila; developmental biology;
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX Drosophila melanogaster.
 OS WO200171042-A2.
 PN 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US09231.
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 DR N-PSDB; ABL14773.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX Disclosure; SEQ ID NO 38802; 21pp + Sequence Listing; English.
 PS The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 204 AA;

Query Match 49.5%; Score 546.5; DB 22; Length 204;
 Best Local Similarity 49.8%; Pred. No. 1.4e-49;
 Matches 105; Conservative 47; Mismatches 50; Indels 9; Gaps 3;
 Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMTIEVDGKIRVI 59
 Db 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMTIEVDGKIRVI 60
 Qy 60 QIWDTAGQERYQTITKQYIRRAQGIPLVDYDSSERSYOHIMKWSDVDYAPGKILIG 119
 Db 61 QIWDTAGQERYQTITKQYIRRAQGIPLVDYDSSERSYOHIMKWSDVDYAPGKILIG 120
 Qy 120 GNKADEQKROVREGQGOOLAKKEYGMDFYETSACTNLNLIKESFTR 179
 Db 121 GNKADEQKROVREGQGOOLAKKEYGMDFYETSACTNLNLIKESFTR 176
 Qy 180 LRMSANALALAELEEKEGKPGPANSKTC 210
 Db 177 RESAENQERVIIIDRRNQEKAP----GYSKCC 203

RESULT 5
 ABB11916
 ID ABB11916 standard; peptide; 213 AA.
 XX
 AC ABB11916;

XX 11-JAN-2002 (first entry)
XX Human rab8 homologue, SEQ ID NO:2286.
XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antiulcer.
XX Homo sapiens.
OS WO200157188-A2.
PN 09-AUG-2001.
XX 05-FEB-2001; 2001WO-US03800.
XX 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-457740/49.
DR N-PSDB; ABA09160.
XX Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
XX e.g. arthritis and cancer -
PS Claim 20; Page 276; 1963pp; English.
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g. myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to

CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.
XX
SQ Sequence 213 AA;
Query Match 49.3%; Score 545; DB 22; Length 213;
Best Local Similarity 52.9%; Pred. No. 2.1e-49;
Matches 99; Conservative 45; Mismatches 4; Indels 0; Gaps 0;
QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFSSHSHTIGVDYDKVRIQ 60
DB 7 MAKTYDYLKLLIGDSGVGKTCVLFSEDAFNSTFTIGIDFKTIFELDKRKLQ 66
QY 61 IWDYAGQRYQTITKQYRRAGQFLVYDYSSEYQHIMKWSVDYFAEGVOKILIG 120
DB 67 IWDYAGQRYQTITKQYRRAGQFLVYDYSSEYQHIMKWSVDYFAEGVOKILIG 126
QY 121 NKADDEQRQYQVREGQQLAKYEGMDFYETACTNLNKESTRLTELVLQAHKKEGL 180
DB 127 NKCDVNDKQYQSKERGENLALDYGIKFMETSAKINVENAFFTLARDIKAKMDKLEGN 186
QY 181 RMRASNE 187
DB 187 SPOGSNQ 193
RESULT 6
ABP41333
ID ABP41333 standard; Protein: 221 AA.
XX AC ABP41333;
XX DT 23-AUG-2002 (first entry)
XX Human ovarian antigen HCGMA67, SEQ ID NO:2465.
XX Human; ovarian antigen; ovary; breast; cancer; tumour;
KW ovarian cancer; breast cancer; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive.
XX Homo sapiens.
OS WO200200677-A1.
XX PN 03-JAN-2002.
XX PD 07-JUN-2001; 2001WO-US18569.
XX PF 07-JUN-2000; 2000US-209467P.
XX PR (HUMA-) HUMAN GENOME SCI INC.
PA Birse CE, Rosen CA;
XX WPI; 2002-147878/19.
XX N-PSDB; ABQ54410.
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g.
PT ovarian cancer), immune disorders, cardiovascular disorders and
PT neurological diseases -

XX		Claim 11; SEQ ID No 2465; 2922pp; English.
PS		The invention relates to 2175 novel human ovarian antigens (ABP41054-ABP4328) and to cDNAs encoding them (ABQ54131-ABQ56305), and also encompasses polypeptides 90% identical and polynucleotides 95% identical to the sequences of the invention. The invention additionally relates to recombinant vectors and host cells comprising human ovarian antigen polynucleotides, antibodies against human ovarian antigens, and the use of ovarian antigen polynucleotides and polypeptides in diagnosing, treating, prognosing or preventing various ovary and/or breast-related disorders. Such conditions include ovarian cancer and breast cancer, metastatic tumours of ovarian or breast origin, reproductive system disorders (e.g., infertility, disorders of pregnancy, anovulation, polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and vaginitis), immune disorders (e.g., congenital and acquired immunodeficiencies, autoimmune ophthalmia), systemic lupus erythematosus) blood-related disorders (e.g., anaemia), cardiovascular disorders, respiratory disorders, neurological disorders, gastrointestinal disorders and urinary system disorders. Ovarian antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate ovarian antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as food additives or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. The present sequence represents a human ovarian antigen of the invention.
CC		Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
CC		
XX		
SQ	Sequence	221 AA;
	Query Match	48.9%; Score 540; DB 23; Length 221;
	Best Local Similarity	52.7%; Pred. No. 7.5e-49;
	Matches	98; Conservative 45; Mismatches 43; Indels 0; Gaps 0
QY	2 AKQVDVFLRLLLIGDSGVGKTCCLCRFTFDNHFHSHISTIGVDFKMKITVEDGIVRIQI	61
DB	16 AKTYDLFKLLLLIGDSGVGKTCVLFRSEDAFNSTFISITGIDFKIRIELDGKRKIQI	75
QY	62 WDTAGQRYSQTITKOYYRRAGQIFELVDISSERSYQHIMKWSDVEYAPEGVKILIGN	121
DB	76 WDTAGQERFRTITATYRGAMGIMLYDITNEKSPDNRNWIRNTEEHASADVERKMILGN	135
QY	122 KADEQRKVREGQQQLAKHYGMDFVETSACTNLNIKESFRTLTELVLQAHKKELEGR	181
DB	136 KCDVNDKRVQSKERGEKLALDYGIKFMTSAKANINVENAFITLARDIKAKMDKKLEGNS	195
QY	182 MRASNE	187
DB	196 PQGSNQ	201
RESULT 7		
AAB58196		
ID	AAB58196 standard; Protein; 246 AA.	
AC	AAB58196;	
XX		
DT	14-MAR-2001 (first entry)	
XX		
DE	Lung cancer associated polypeptide sequence SEQ ID 534.	
KW	Human; lung cancer associated protein; neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular active; vulnerary;	
KW	gastrointestinal; nephrotropic; antiinfective; gynecological;	
KW	antibacterial; diagnosis; neural disorder; immune disorder; reproductive;	
KW	proliferative disorder; wound healing; infectious disease.	
OS	Homo sapiens.	

DE Amino acid sequence of human RAB10.

KW RAB protein; GTPase; GTP binding; gene therapy; cancer.

XX Homo sapiens.

XX WO2000058464-A2.

XX PD 05-OCT-2000.

XX PF 13-MAR-2000; 2000WO-US06330.

XX PR 25-MAR-1999; 99US-0126083.

XX PA (AXYS-) AXYS PHARM INC.

XX PI Allen M, Abel K, McIntosh B, Vega R, Rutter M, Buckler AJ;

XX DR WPI; 2000-647233/62.

XX DR N-PSDB; AAA96887.

XX PT Novel isolated nucleic acid encoding a mammalian RAB protein useful for identifying homologous or related genes, in producing composition that modulates expression or function of RAB for cancer therapy.

XX PS Claim 2; Page 39; 58pp; English.

XX CC The present sequence represents a mammalian RAB protein. RAB proteins constitute the largest family of small GTPases, with over 40 identified isoforms. RAB proteins contain four highly conserved peptide sequences involved in GTP binding and hydrolysis. Compositions comprising RAB nucleic acid are useful for identifying homologous or related genes, in producing compositions that modulate the expression or function of RAB, for gene therapy, mapping functional regions of the protein and in studying associated physiological pathways. In addition, modulation of the gene activity in vivo is used for prophylactic and therapeutic purposes, such as treatment of cancer, and identification of cell type based on expression. The DNA may also be used to identify expression of the gene in a biological specimen.

XX SQ Sequence 200 AA;

Query Match 48.0%; Score 530.5; DB 21; Length 200;

Best Local Similarity 55.8%; Pred. No. 6.6e-48;

Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;

QY 1 MAKO-YDVLFRLLIGDSGVGKTCILCRFTDNEFHSHSHSTIGVDFKMTIEVDGIKVI 59

Db 1 MAKTYDLLFKLLIGDSGVGKTCVLFPSDDAFTTFTISGIDFKIKTVLQGGKIKL 60

QY 60 QIWDTAGOERTQITKQYRRAQGIPLYDISSERSYOHIMKWSVDVEYAPGVOKILI 119

Db 61 QIWDTAGOERTHTTTSYRGAMGIMLYDITNGKSFENISKWLNIDEHANEDVERMLL 120

QY 120 GNKADEOKRGVREGOOLAKKEYGMDFYETSACTNLNKKSEFTLTELVLQ 171

Db 121 GNKCDMDKRVVPKGEQIAREHGIRFETSANINIERAKFTLAEDILR 172

RESULT 9

AAB09979

XX ID AAB09979 standard; Protein; 200 AA.

XX AC AAB09979;

XX DT 19-OCT-2000 (first entry)

XX DE Human RAB10 protein.

XX KW Human; RAB10; Yap/Rab family.

XX OS Homo sapiens.

XX PA (HELI-) HELIX RES INST.

PN CN1249345-A.

XX PD 05-APR-2000.

XX PF 28-SEP-1998; 98CN-0121911.

XX PR 28-SEP-1998; 98CN-0121911.

XX PA (UYFU-) UNIV FUDAN.

XX PI Yu L, Tu Q, Gao J;

XX DR WPI; 2000-400723/35.

XX DR N-PSDB; AAA40104.

XX PT Preparation of human gene coding sequence, its encoded polypeptide -

XX PS Claim 2; Page 18; 23pp; Chinese.

XX CC This invention describes a novel human RAB10 cDNA sequence. The protein coded by said sequence belongs to Yap/Rab family and is the homolog of mouse RAB10. The present invention also relates to the polypeptide coded by said nucleotide sequence and the application and preparing process of said polynucleotide and said polypeptide. This sequence represents the human RAB10 protein described in the method of the invention.

XX SQ Sequence 200 AA;

Query Match 48.0%; Score 530.5; DB 21; Length 200;

Best Local Similarity 55.8%; Pred. No. 6.6e-48;

Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;

QY 1 MAKO-YDVLFRLLIGDSGVGKTCILCRFTDNEFHSHSHSTIGVDFKMTIEVDGIKVI 59

Db 1 MAKTYDLLFKLLIGDSGVGKTCVLFPSDDAFTTFTISGIDFKIKTVLQGGKIKL 60

QY 60 QIWDTAGOERTQITKQYRRAQGIPLYDISSERSYOHIMKWSVDVEYAPGVOKILI 119

Db 61 QIWDTAGOERTHTTTSYRGAMGIMLYDITNGKSFENISKWLNIDEHANEDVERMLL 120

QY 120 GNKADEOKRGVREGOOLAKKEYGMDFYETSACTNLNKKSEFTLTELVLQ 171

Db 121 GNKCDMDKRVVPKGEQIAREHGIRFETSANINIERAKFTLAEDILR 172

RESULT 10

AAB95340

XX ID AAB95340 standard; Protein; 200 AA.

XX AC AAB95340;

XX DT 26-JUN-2001 (first entry)

XX DE Human protein sequence SEQ ID NO:17619.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

XX PA (HELI-) HELIX RES INST.

PD 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT;
 PI WPI; 2001-639362/73.
 XX N-PSDB; AAS71453.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID NO 37625; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 218 AA;
 Query Match 47.9%; Score 529.5; DB 22; Length 218;
 Best Local Similarity 55.8%; Pred. No. 9.6e-48;
 Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;
 QY 1 MAKO-YDVLFRLLIGDSGVGKTCILCRFTDNEFHSHISTIGVDFKMTIEVDGIKVI 59
 DB 19 MAKTYDLFLKLLIGDSGVGKTCVLFPSDAENFTFSTIGIDFKIKTVELQGGKIKL 78
 QY 60 QIWDTAGOERYQTITKQYRRAGGFLVYDYSERSYOHIMKWSVDVEYAPGVOKILI 119
 DB 79 QIWDTAGOERHTTITSYRGAMGLVYDITNGKSPENISKWLINDEHANEDVERML 138
 QY 120 GNKADEQKRVGREGQQLAKKEYGMDFYETSACTNLNIKESFTRLTELVLQ 171
 DB 139 GNKCDMDKRVVPKRGQIAREHGIRFETSANANIEKAFLLTAEILR 190
 RESULT 13
 ABB71647
 ID ABB71647 standard; Protein: 207 AA.
 XX ABB71647;
 AC ABB71647;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 41733.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical.
 XX Drosophila melanogaster.
 XX WO200171042-A2.
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US09231.
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 XX N-PSDB; ABL15750.
 DR
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions
 XX
 PS Disclosure; SEQ ID NO 41733; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 207 AA;
 Query Match 47.8%; Score 528.5; DB 22; Length 207;
 Best Local Similarity 53.8%; Pred. No. 1.1e-47;
 Matches 100; Conservative 41; Mismatches 40; Indels 5; Gaps 1;
 QY 1 MAKQYDVLFRLLIGDSGVGKTCILCRFTDNEFHSHISTIGVDFKMTIEVDGIKVI 60
 DB 1 MAKTYDYLKLLIGDSGVGKTCILFRSEDAFNTFTSTIGIDFKIKTIELDNKKIKLQ 60
 QY 61 IWDTAGOERYQTITKQYRRAGGFLVYDYSERSYOHIMKWSVDVEYAPGVOKILIG 120
 DB 61 IWDTAGOERFTTITAYYRGAMGLVYDITNGKSPENIKWIRNIEENASADVKEKMLG 120
 QY 121 NKADDEQKRVGREGQQLAKKEYGMDFYETSACTNLNIKESFTRLTELVLQAHKKEGL 180
 DB 121 NKCELTDRQVSKERGEQLAIEYGIKFMETSAKASINVERAF-----LTASDIKAKTEK 175
 QY 181 RMRASN 186
 DB 176 RMEANN 181
 RESULT 14
 AAG67154
 ID AAG67154 standard; Protein: 207 AA.
 XX AAG67154;
 AC AAG67154;
 XX 13-NOV-2001 (first entry)
 DT
 XX Amino acid sequence of human 27423 G-protein.
 DE Human; G-protein; 32705; 23224; 27423; 32700; 32712; lung disorder;
 KW congenital anomaly; pulmonary congestion; oedema; haemorrhage;

KW adult respiratory distress syndrome; Goodpasture's syndrome;
 KW chronic obstructive pulmonary disease; aschma; pulmonary hypertension;
 KW liver disorder; hepatic injury; jaundice; cholestasis; viral hepatitis;
 KW cirrhosis; Wilson's disease; autoimmune hepatitis; hepatic failure;
 KW brain disorder; hypoxia; cerebral ischemia; intracranial haemorrhage;
 KW acute meningitis; Parkinson's disease; Alzheimer's disease; glioma;
 KW chronic bacterial meningencephalitis; multiple sclerosis;
 KW ankyrotic lateral sclerosis; stroke; Huntington's disease.
 XX
 OS Homo sapiens.
 XX
 PN W0200164887-A2.
 XX
 PD 07-SEP-2001.
 XX
 XX 27-FEB-2001; 2001WO-US06292.
 PF
 XX 29-FEB-2000; 2000US-0185606.
 PR
 XX (MILL-) MILLENNIUM PHARM INC.
 PA
 PI Meyers RA;
 XX
 XX WPI: 2001-550182/61.
 DR N-PSDB; AAG87154.
 DR
 XX Novel human small G-protein polypeptides and polynucleotides for
 PT treating lung disorders, liver disorders and brain disorders -
 PT
 XX Claim 8; Fig 15; 151pp; English.
 PS
 XX

CC The present sequence represents a human G-protein. The specification
 CC describes 32705, 23224, 27423, 32700 or 32712 small G-proteins. The
 CC G-protein polypeptides and polynucleotides are useful as a target for
 CC diagnosis and treatment of G-protein mediated or related disorders,
 CC and for identifying agonists and antagonists for diagnosis and
 CC treatment. They are useful for treating disorders of lung (e.g.
 CC congenital anomalies, pulmonary congestion, oedema, adult respiratory
 CC distress syndrome, haemorrhage, chronic obstructive pulmonary disease,
 CC asthma, Goodpasture's syndrome and pulmonary hypertension), liver
 CC (e.g. hepatic injury, jaundice, cholestasis, viral hepatitis, cirrhosis,
 CC Wilson's disease, autoimmune hepatitis and hepatic failure), and
 CC brain (e.g. hypoxia, cerebral ischemia, intracranial haemorrhage, acute
 CC meningitis, Parkinson's disease, Alzheimer's disease, gliomas, chronic
 CC bacterial meningencephalitis, multiple sclerosis, amyotrophic lateral
 CC sclerosis, stroke and Huntington's disease).
 XX
 SQ Sequence 207 AA;

Query Match 47.8%; Score 528.5; DB 22; Length 207;
 Best Local Similarity 54.0%; Pred. No. 1.1e-47;
 Matches 95; Conservative 47; Mismatches 33; Indels 1; Gaps 1;
 QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHISTIGVDKMTIEVDGKVRQ 60
 DB 1 MARYDYLFKLLIGDSGVGKTCCLFRFSDAFTFTSTIGDKTIELDGKIKLQ 60
 QY 61 IWDTAGQERYQTITKQYRRAQGIYFLYDTSERSYQHIMKWSVDVEYAPGKILIG 120
 DB 61 IWDTAGQERYQTITAYRGAMGIMLYDITNEKSPDNKWNIRNIEHASSDVERMILG 120
 QY 121 NKADEQKRVGREGQOOLAKEYGMDFYETSACNLNIKES-FTRLVELVLAHRK 175
 DB 121 NKCDMDKRVQSKERGEKLAIDYGIKFLTSKSSANVEEAFETLARDIMTKLNK 176

RESULT 15
 AAB92628
 ID AAB92628 standard; Protein; 207 AA.

XX
 AC AAB92628;
 XX
 XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:10930.
 DE
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 KW
 XX Homo sapiens.
 OS
 XX EP1074617-A2.
 PN
 XX 07-FEB-2001.
 PD
 XX 28-JUL-2000; 2000EP-0116126.
 PF
 XX 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 PR
 XX (HELI-) HELIX RES INST.
 PA
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 PI
 XX WPI: 2001-318749/34.
 DR
 XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 PT
 XX Claim 8; SEQ ID 10930; 2537pp + CD ROM; English.
 PS
 XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 207 AA;

Query Match 47.8%; Score 528.5; DB 22; Length 207;
 Best Local Similarity 54.0%; Pred. No. 1.1e-47;
 Matches 95; Conservative 47; Mismatches 33; Indels 1; Gaps 1;
 QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHISTIGVDKMTIEVDGKVRQ 60
 DB 1 MARYDYLFKLLIGDSGVGKTCCLFRFSDAFTFTSTIGDKTIELDGKIKLQ 60
 QY 61 IWDTAGQERYQTITKQYRRAQGIYFLYDTSERSYQHIMKWSVDVEYAPGKILIG 120
 DB 61 IWDTAGQERYQTITAYRGAMGIMLYDITNEKSPDNKWNIRNIEHASSDVERMILG 120
 QY 121 NKADEQKRVGREGQOOLAKEYGMDFYETSACNLNIKES-FTRLVELVLAHRK 175
 DB 121 NKCDMDKRVQSKERGEKLAIDYGIKFLTSKSSANVEEAFETLARDIMTKLNK 176

Search completed: June 18, 2003, 15:33:30
Job time : 72 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:30:31 ; Search time 40 Seconds
(without alignments)
509.512 Million cell updates/sec

Title: US-09-817-198B-2
Perfect score: 1105
Sequence: 1 MAKQYDVLFRLLIGDSGVG.....LEEEGKPEGPANSSKTCWC 212

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1077	97.5	212	2 F42148	GTP-binding protei
2	559	50.6	224	2 T33855	hypothetical prote
3	545	49.3	207	2 B49647	GTP-binding protei
4	545	49.3	207	2 B36364	GTP-binding protei
5	532	48.1	203	2 A49647	GTP-binding protei
6	531.5	48.1	200	2 A38625	GTP-binding protei
7	530.5	48.0	200	2 D36364	GTP-binding protei
8	530	48.0	209	2 B38625	GTP-binding protei
9	526	47.6	206	2 T78851	GTP-binding protei
10	522.5	47.3	200	2 B42148	GTP-binding protei
11	519	47.0	216	2 T45901	GTPase ATRAB8 - Ar
12	514	46.5	201	2 T38971	hypothetical prote
13	512	46.3	216	2 S33900	GTP-binding protei
14	512	46.3	216	2 T48378	GTP-binding protei
15	510	46.2	200	2 S12790	GTP-binding protei
16	509	46.1	215	2 T14565	GTP-binding protei
17	508.5	46.0	215	2 S7478	GTP-binding protei
18	508	46.0	216	2 J30640	GTP-binding protei
19	507.5	45.9	222	2 T14405	small GTP-binding
20	506	45.8	216	2 S57471	GTP-binding protei
21	505.5	45.7	217	2 S36365	GTP-binding protei
22	502.5	45.5	215	2 S7462	GTP-binding protei
23	501	45.3	203	2 S14935	GTP-binding protei
24	495.5	44.8	204	2 J37589	Sec4p homolog - ye
25	492.5	44.6	215	2 S57474	GTP-binding protei
26	483.5	43.8	208	2 A34716	GTP-binding protei
27	483	43.7	208	2 A38202	GTP-binding protei
28	482	43.6	203	2 S34253	GTP-binding protei
29	480	43.4	202	2 S38740	GTP-binding protei

30	479	43.3	203	2 B34716	GTP-binding protei
31	478	43.3	202	2 J72515	GTP-binding protei
32	477.5	43.2	201	2 J2488	GTP-binding protei
33	477.5	43.2	203	2 J4105	GTP-binding protei
34	477	43.2	201	2 J5337	GTP-binding protei
35	477	43.2	202	2 S41430	GTP-binding protei
36	476	43.1	258	2 B86153	ARA-5 [imported] -
37	475.5	43.0	203	2 J1247	ARA-5 [imported] -
38	475	43.0	203	2 S30096	GTP-binding protei
39	473.5	42.9	205	2 T33781	hypothetical prote
40	473	42.8	206	2 T14391	GTP-binding protei
41	473	42.8	210	2 T18242	ras protein homolo
42	470.5	42.6	203	2 B38202	GTP-binding protei
43	470.5	42.6	215	1 TVBYQ4	GTP-binding protei
44	469	42.4	205	2 S38339	GTP-binding protei
45	468	42.4	205	1 TVRTYP	GTP-binding protei

ALIGNMENTS

RESULT 1

F42148
GTP-binding protein rab15 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 02-Feb-2001
C:Accession: F42148
R:Elferink, L.A.; Anzai, K.; Scheller, R.H.
J. Biol. Chem. 267, 5768-5775, 1992
A:Title: rab15, a novel low molecular weight GTP-binding protein specifically express
A:Reference number: A42148; MUID:92210533; PMID:1313420
A:Accession: F42148
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-212 <ELF>
A:Cross-references: GB:M33679; NID:9206536; PIDN:AAA41995.1; PID:9206537
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: GTP binding; lipoprotein; nucleotide binding; P-loop; prenylated cysteine
F:9-124/Domain: translation elongation factor Tu homology <ETU>
F:15-22/Region: nucleotide-binding motif A (P-loop)
F:121-124/Region: GTP-binding NKXD motif
F:210,212/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match	97.5%	Score 1077;	DB 2;	Length 212;
Best Local Similarity	97.6%	Pred. NO. 9.2e-79;		
Matches 207;	Conservative 2;	Mismatches 3;	Indels 0;	Gaps 0;
QY	1	MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHSTIGVDFKMTIEVDGKVRIQ	60	
Db	1	MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHSTIGVDFKMTIEVDGKVRIQ	60	
QY	61	IWDTAGQERYQTIKQYRRAQGIPLVYDIDISERSYQHIMKWSDVDEYAPGVOKILIG	120	
Db	61	IWDTAGQERYQTIKQYRRAQGIPLVYDIDISERSYQHIMKWSDVDEYAPGVOKILIG	120	
QY	121	NKADEQKRVGREGOQOLAKEYGMDFYETSACTNLNIKESFTRTELVLQAHRKEGL	180	
Db	121	NKADEQKRVGREGOQOLAKEYGMDFYETSACTNLNIKESFTRTELVLQAHRKEGL	180	
QY	181	RRASNELALAELEEEGKPEGPANSSKTCWC	212	
Db	181	RTCASNELALAELEDEGKTEGPANSSKTCWC	212	

RESULT 2

T33855
hypothetical protein D1037.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 11-Jan-2000
C:Accession: T33855
R:Ledwith, J.; Biewald, T.
Submitted to the EMBL Data Library, November 1998
A:Description: The sequence of C. elegans cosmid D1037.

C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop;
F;9-124/Domain: translation elongation factor Tu homology <ETU>
F;15-22/Region: nucleotide-binding motif A (P-loop)
F;121-124/Region: GTP-binding NAKD motif
F;151-153/Region: GTP-binding NAKD motif
F;21,22,40,121,122,124,151/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser) #sta
F;200/Binding site: farnesyl (Cys) (covalent) #status predicted

Query Match 48.1%; Score 532; DB 2; Length 203;
Best Local Similarity 47.6%; Pred. No. 2.6e-35;
Matches 100; Conservative 49; Mismatches 43; Indels 18; Gaps 3;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDFKMTIEVDGKVRQ 60
DQ 1 MAKAYDHLFKLLIGDSGVGKTCCLIRFAEDNNTYISTIGIDFKIRTVDEGKKIKLQ 60
QY 61 IWDTAGQERYQTITKQYRRAQGIPLVYDIDSSERSYQHIMKWSVDVDEYAPGKILIG 120
DQ 61 IWDTAGQERYQTITKQYRRAQGIPLVYDIDSSERSYQHIMKWSVDVDEYAPGKILIG 120
QY 121 NKADEQKROVGRGQOQLAKYGMDFYETSACTNLNIKESFTRLTFLVLAQHRKEGL 180
DQ 121 NKCDMEAKRVQEQADKLAREHIGIRFFETSAKSMNVNDEAFSSLDAIL-----LKSG 174
QY 181 RMRASNELALAELEEGKPEGANSSKTC 210
DQ 175 GRNSGN-----GNKP--PSTDLKTC 192

RESULT 6
A38625
GTP-binding protein oral - electric ray (Discopyge ommata)
C;Species: Discopyge ommata
C;Date: 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change 02-Feb-2001
R;Ngsee, J.K.; Elferink, L.A.; Scheller, R.H.
J. Biol. Chem. 266, 2675-2680, 1991
A;Title: A family of ras-like GTP-binding proteins expressed in electromotor neurons.
A;Reference number: A38625; MUID:91115900; PMID:1899244
A;Accession: A38625
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-200 <NGS>
A;Cross-references: GB:X38390; NID:g213114; PIDN:AAA49230.1; PID:g213115
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; membrane trafficking; nucleotide binding; P-loop
F;10-125/Domain: translation elongation factor Tu homology <ETU>
F;16-23/Region: nucleotide-binding motif A (P-loop)
F;122-125/Region: GTP-binding NAKD motif
F;152-154/Region: GTP-binding NAKD motif

Query Match 48.1%; Score 531.5; DB 2; Length 200;
Best Local Similarity 56.4%; Pred. No. 2.8e-35;
Matches 97; Conservative 44; Mismatches 30; Indels 1; Gaps 1;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDFKMTIEVDGKVRQ 59
DQ 1 MAKAYDHLFKLLIGDSGVGKTCCLIRFAEDNNTYISTIGIDFKIRTVDEGKKIKLQ 60
QY 60 QIWDTAGQERYQTITKQYRRAQGIPLVYDIDSSERSYQHIMKWSVDVDEYAPGKILIG 119
DQ 61 QIWDTAGQERYQTITKQYRRAQGIPLVYDIDSSERSYQHIMKWSVDVDEYAPGKILIG 120
QY 120 GNKADEQKROVGRGQOQLAKYGMDFYETSACTNLNIKESFTRLTFLVLAQHRKEGL 171
DQ 121 GNKCDMDKRVVPRGKEQIAREHIGIRFFETSAKSNINIEKAFITLAEIDILQ 172

RESULT 7
D36364
GTP-binding protein rab10 - dog
C;Species: Canis lupus familiaris (dog)
C;Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 02-Feb-2001

C;Accession: D36364; S15601
R;Chavrier, P.; Vingron, M.; Sander, C.; Simons, K.; Zerial, M.
Mol. Cell. Biol. 10, 6578-6585, 1990
A;Title: Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell line.
A;Reference number: A36364; MUID:91061765; PMID:2123294
A;Accession: D36364
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-200 <CHA>
A;Cross-references: GB:X56387; NID:g914; PIDN:CAA39798.1; PID:g915
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop
F;10-125/Domain: translation elongation factor Tu homology <ETU>
F;16-23/Region: nucleotide-binding motif A (P-loop)
F;122-125/Region: GTP-binding NAKD motif
F;152-154/Region: GTP-binding NAKD motif
F;199,200/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 48.0%; Score 530.5; DB 2; Length 200;
Best Local Similarity 55.8%; Pred. No. 3.3e-35;
Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDFKMTIEVDGKVRQ 59
DQ 1 MAKAYDHLFKLLIGDSGVGKTCCLIRFAEDNNTYISTIGIDFKIRTVDEGKKIKLQ 60
QY 60 QIWDTAGQERYQTITKQYRRAQGIPLVYDIDSSERSYQHIMKWSVDVDEYAPGKILIG 119
DQ 61 QIWDTAGQERYQTITKQYRRAQGIPLVYDIDSSERSYQHIMKWSVDVDEYAPGKILIG 120
QY 120 GNKADEQKROVGRGQOQLAKYGMDFYETSACTNLNIKESFTRLTFLVLAQHRKEGL 171
DQ 121 GNKCDMDKRVVPRGKEQIAREHIGIRFFETSAKSNINIEKAFITLAEIDILQ 172

RESULT 8
B38625
GTP-binding protein ora2 - electric ray (Discopyge ommata)
C;Species: Discopyge ommata
C;Date: 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change 02-Feb-2001
A;Accession: B38625
R;Ngsee, J.K.; Elferink, L.A.; Scheller, R.H.
J. Biol. Chem. 266, 2675-2680, 1991
A;Title: A family of ras-like GTP-binding proteins expressed in electromotor neurons.
A;Reference number: A38625; MUID:91115900; PMID:1899244
A;Accession: B38625
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-209 <NGS>
A;Cross-references: GB:X38391
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; lipoprotein; membrane trafficking; nucleotide binding; P-loop
F;9-124/Domain: translation elongation factor Tu homology <ETU>
F;15-22/Region: nucleotide-binding motif A (P-loop)
F;121-124/Region: GTP-binding NAKD motif
F;150-152/Region: GTP-binding NAKD motif
F;206/Binding site: geranyl-geranyl (Cys) (covalent) #status predicted

Query Match 48.0%; Score 530; DB 2; Length 209;
Best Local Similarity 47.4%; Pred. No. 3.9e-35;
Matches 99; Conservative 56; Mismatches 44; Indels 10; Gaps 3;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHISTIGVDFKMTIEVDGKVRQ 60
DQ 1 MAKAYDHLFKLLIGDSGVGKTCCLIRFAEDNNTYISTIGIDFKIRTVDEGKKIKLQ 60
QY 61 IWDTAGQERYQTITKQYRRAQGIPLVYDIDSSERSYQHIMKWSVDVDEYAPGKILIG 120
DQ 61 IWDTAGQERYQTITKQYRRAQGIPLVYDIDSSERSYQHIMKWSVDVDEYAPGKILIG 120
QY 121 NKADEQKROVGRGQOQLAKYGMDFYETSACTNLNIKESFTRLTFLVLAQHRKEGL 180
DQ 121 NKCDMEAKRVQEQADKLAREHIGIRFFETSAKSNINIEKAFITLAEIDILQ 172

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:24:36 ; Search time 22 Seconds
(without alignments)
399.681 Million cell updates/sec

Title: US-09-817-198b-2

Perfect score: 1105

Sequence: 1 MAKQYDVLFRLLIGSGVG.....LEEEKGPEGPANSTKWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1077	97.5	212	1 RB15_RAT	P35289 rattus norv
2	545	49.3	207	1 RAB8_HUMAN	P24407 homo sapien
3	539.5	48.8	210	1 RAB8_DISOM	P22128 discopoyge o
4	532	48.1	203	1 RB13_HUMAN	P51153 homo sapien
5	531.5	48.1	200	1 RAOL_DISOM	P22127 discopoyge o
6	530.5	48.0	200	1 RB10_CANFA	P24409 canis famil
7	530.5	48.0	200	1 RB10_HUMAN	O88386 homo sapien
8	528.5	47.8	207	1 RB8B_HUMAN	Q92930 homo sapien
9	528.5	47.8	207	1 RB8B_RAT	P70530 rattus norv
10	526	47.6	206	1 RAB8_MOUSE	P55238 mus musculu
11	510	46.2	200	1 YPT2_SCHPO	P17609 schizosacch
12	509	46.1	215	1 RAB1_BETVU	Q39433 beta vulgar
13	508	46.0	216	1 ARA3_ARATH	P28186 arabidopsis
14	505.5	45.7	217	1 YPT2_VOLCA	P36861 volvox cart
15	502.5	45.5	200	1 RB10_RAT	P35281 rattus norv
16	502	45.4	203	1 RYLL_YARLI	P41924 yarrowia li
17	483.5	43.8	208	1 SAS1_DICDI	P20790 dictyosteli
18	483	43.7	208	1 YPT1_MAZE	P16976 zea mays (m
19	480	43.4	202	1 RIC1_ORYSA	P40392 oryza sativ
20	479	43.3	203	1 SAS2_DICDI	P20791 dictyosteli
21	477.5	43.2	201	1 RB35_HUMAN	Q15286 homo sapien
22	477.5	43.2	203	1 YPT1_CHLRE	Q39571 chlamydomon
23	477	43.2	201	1 YPT1_PHYIN	Q01890 phytophthor
24	476	43.1	258	1 ARA5_ARATH	P28188 arabidopsis
25	475.5	43.0	203	1 YPT1_VOLCA	P31584 volvox cart
26	475	43.0	203	1 YPT1_NEUCR	P33723 neurospora
27	473	42.8	210	1 SEC4_CANAL	Q14462 candida alb
28	470.5	42.6	203	1 YPT2_MAZE	Q05737 zea mays (m
29	470.5	42.6	215	1 SEC4_YEAST	P07560 saccharomyc
30	469	42.4	201	1 RB1B_RAT	P10536 rattus norv
31	469	42.4	205	1 RAB1_LYMST	Q05974 lymnaea sta
32	468	42.4	205	1 RB1A_RAT	P05711 rattus norv
33	467	42.3	203	1 YPT1_SCHPO	P11620 schizosacch

ALIGNMENTS

RESULT 1

ID	RB15_RAT	STANDARD;	PRT;	212 AA.
AC	P35289;			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	Ras-related protein Rab-15.			
GN	RAB15.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
[1]				
RP	SEQUENCE FROM N.A.			
RC	STRAIN-Sprague-Dawley; TISSUE-Brain;			
RX	MEDLINE=92210533; PubMed=1313420;			
RA	Elferink L.A., Anzai K., Scheller R.H.;			
RT	"Rab15, a novel low molecular weight GTP-binding protein specifically			
RT	expressed in rat brain.";			
RL	J. Biol. Chem. 267:5768-5775(1992).			
[2]				
RP	ERRATUM.			
RX	MEDLINE=93054572; PubMed=1429617;			
RA	Elferink L.A., Anzai K., Scheller R.H.;			
RL	J. Biol. Chem. 267:22693-22693(1992).			
CC	-1- FUNCTION: MAY ACT IN CONCERT WITH RAB3A IN REGULATING ASPECTS OF			
CC	SYNAPTIC VESICLE MEMBRANE FLOW WITHIN THE NERVE TERMINAL.			
CC	-1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN NEURAL TISSUES.			
CC	-1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; M83679; AAA41995.1; -			
DR	PIR; F42148; F42148.			
DR	HSSP; P05713; 3RAB.			
DR	InterPro; IPR003579; GTPase_Rab.			
DR	InterPro; IPR001806; Ras_trnsfrmng.			
DR	InterPro; IPR005225; Small_GTP.			
DR	Pfam; PF00071; ras; 1.			
DR	PRINTS; PR00449; RASTRNSFRMNG.			
DR	SMART; SM00175; RAB; 1.			
DR	TIGRFAMS; TIGR00231; small_GTP; 1.			
KW	GTP-binding; Lipoprotein; Prenylation; protein transport.			
FT	NP_BIND 15 22 GTP (BY SIMILARITY).			
FT	NP_BIND 63 67 GTP (BY SIMILARITY).			
FT	NP_BIND 121 124 GTP (BY SIMILARITY).			
FT	LIPID 210 210 GERANYL-GERANYL (BY SIMILARITY).			
FT	LIPID 212 212 GERANYL-GERANYL (BY SIMILARITY).			
FT	SEQUENCE 212 AA; 24283 MW; 04817DDA66CADE12 CRC64;			
SQ				

P11476 homo sapien
P22125 discopoyge o
P01123 saccharomyc
Q63482 rattus norv
Q9ulw5 homo sapien
P51156 rattus norv
P25228 drosophila
O95716 homo sapien
P11023 bos taurus
O63942 rattus norv
P05713 mus musculu
P20336 homo sapien

```

Query Match      97.5%; Score 1077; DB 1; Length 212;
Best Local Similarity 97.6%; Pred. No. 6.9e-80;
Matches 207; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSTIGVDFPKMTIEVDGKVRQ 60
DB 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSTIGVDFPKMTIEVDGKVRQ 60

QY 61 IWDTAGQRYQTITKQYRRAGIFLVYDISERSYQHIMKWSVDVEYAPGVOKILIG 120
DB 61 IWDTAGQRYQTITKQYRRAGIFLVYDISERSYQHIMKWSVDVEYAPGVOKILIG 120

QY 121 NKADEQRQVREGOQOLAKYGMDFYETSACTNLIKESFTRTLVLQAHKREGL 180
DB 121 NKADEQRQVREGOQOLAKYGMDFYETSACTNLIKESFTRTLVLQAHKREGL 180

QY 181 RMRASNELALAELEEGKTEGPANSSKTCWC 212
DB 181 RMRASNELALAELEEGKTEGPANSSKTCWC 212

RESULT 2
RAB8_HUMAN
ID RAB8_HUMAN STANDARD; PRT; 207 AA.
AC P24407;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Ras-related protein Rab-8 (Oncogene c-mel).
GN RAB8 OR MEL.
OS Homo sapiens (Human), and
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606, 9615;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-Human;
RX MEDLINE=94124602; PubMed=8294494;
RA Zahraoui A., Joberty G., Arpin M., Fontaine J.J., Hellio R.,
RA Tavittian A., Louvard D.;
RT "A small rab GTPase is distributed in cytoplasmic vesicles in non
RT polarized cells but colocalizes with the tight junction marker ZO-1
RT in polarized epithelial cells.";
RL J. Cell Biol. 124:101-115(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES-Human;
RX MEDLINE=91360267; PubMed=1886711;
RA Nimmo E.R., Sanders P.G., Padua R.A., Hughes D., Williamson R.,
RA Johnson K.J.;
RT "The MEL gene: a new member of the RAB/YPT class of RAS-related
RT genes.";
RL Oncogene 6:1347-1351(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES-Human;
RX Strausberg R.;
RA Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES-C.familiaris; STRAIN-Cocker spaniel;
RX MEDLINE=91061765; PubMed=2123294;
RA Chavrier P., Vingron M., Sander C., Simons K., Zerial M.;
RT "Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell
RT line.";
RL Mol. Cell. Biol. 10:6578-6585(1990).
CC -!- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
CC NEUROTRANSMITTER RELEASE.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X56741; CAA40065.1; -
DR EMBL; X53268; AAB19681.1; -
DR EMBL; BC002977; AAH02977.1; -
DR EMBL; X56385; CAB56776.1; -
DR PIR; B36364; B36364.
DR PIR; B49647; B49647.
DR HSSP; P05713; 3RAB.
DR GENE; HGNC:7007; MEL.
DR MIM; 165040; -
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001230; Prenyl_site.
DR InterPro; IPR001806; Ras_transfmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFS; TIGR00231; small_GTP; 1.
KW GTP-binding; Prenylation; Lipoprotein; Protein transport;
KW Proto-oncogene.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 204 204 GERANYL-GERANYL (BY SIMILARITY).
FT CONFLICT 177 183 LEGNSPQ -> WKATAP (IN REF. 2).
SQ SEQUENCE 207 AA; 23668 MW; AA52DBF54A2CD056 CRC64;

Query Match      49.3%; Score 545; DB 1; Length 207;
Best Local Similarity 52.9%; Pred. No. 4.2e-37;
Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSTIGVDFPKMTIEVDGKVRQ 60
DB 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSTIGVDFPKMTIEVDGKVRQ 60

QY 61 IWDTAGQRYQTITKQYRRAGIFLVYDISERSYQHIMKWSVDVEYAPGVOKILIG 120
DB 61 IWDTAGQRYQTITKQYRRAGIFLVYDISERSYQHIMKWSVDVEYAPGVOKILIG 120

QY 121 NKADEQRQVREGOQOLAKYGMDFYETSACTNLIKESFTRTLVLQAHKREGL 180
DB 121 NKADEQRQVREGOQOLAKYGMDFYETSACTNLIKESFTRTLVLQAHKREGL 180

QY 181 RMRASNE 187
DB 181 SPQGSNQ 187

RESULT 3
RAB8_DISOM
ID RAB8_DISOM STANDARD; PRT; 210 AA.
AC P22128;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Ras-related protein Rab-8 (ORF2).
OS Discopige omata (Electric ray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Hypnosqualea; Pristiorajae; Batoidae;
OC Torpediniformes; Narcinidae; Narcinidae; Discopige.
OX NCBI_TaxID=7785;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Electric lobe;
RX MEDLINE=91115900; PubMed=1899244;
RA Ngsee J.K., Elferink L.A., Scheller R.H.;

```

```

RT "A family of ras-like GTP-binding proteins expressed in electromotor
RL J. Biol. Chem. 266:2675-2680(1991).
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M38391; AAA49232.1; --
CC PIR; B38625; B38625.
CC HSSP; P05713; 3RAB.
CC InterPro; IPR003579; GTPase_Rab.
CC InterPro; IPR001230; Prenyl_site.
CC InterPro; IPR001806; Ras_trnsmg.
CC InterPro; IPR005225; Small_GTP.
CC Pfam; PF00071; Ras; 1.
CC PRINTS; PR00449; RASTRNSFRMNG.
CC SMART; SM00175; RAB; 1.
CC TIGREMS; TIGR00231; small_gtp; 1.
CC GTP-binding; Prenylation; Lipoprotein.
CC NP_BIND 15 22 GTP (BY SIMILARITY).
CC NP_BIND 63 67 GTP (BY SIMILARITY).
CC NP_BIND 121 124 GTP (BY SIMILARITY).
CC DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
CC LIPID 207 207 GERANYL-GERANYL (BY SIMILARITY).
CC SEQUENCE 210 AA; 24164 MW; 0F73EDB0DB9BEA CRC64;

Query Match 48.8%; Score 539.5; DB 1; Length 210;
Best Local Similarity 47.4%; Pred. No. 1.2e-36;
Matches 99; Conservative 56; Mismatches 45; Indels 9; Gaps 2;

Qy 1 MAKQYDLVFLRLIGDSGVGKTCLLCRFTDNEFHSHISTIGVDFKMTIEVDGKVIQ 60
Db 1 MAKTYDLVFLRLIGDSGVGKTCLLCRFTDNEFHSHISTIGVDFKMTIEVDGKVIQ 60
Qy 61 IWDTAGQERYQTITKQYRRAQIGFLVYDISERSYQHIMKWSVDVEYAPGQKILIG 120
Db 61 IWDTAGQERYQTITKQYRRAQIGFLVYDISERSYQHIMKWSVDVEYAPGQKILIG 120
Qy 121 NKADEQKRVGREGQOOLAKEYGMDFYETSACTNLIKESFTRLTSLVLAQHRKEGL 180
Db 121 NKCDMEKRVQSKERGLAIDYGIKEFSAKSSINVEAFITLARDIMTKLNM--- 177
Qy 181 RMRASNELALAE-LEEKEGKPEGPANSSK 208
Db 178 -----NENSLQEAVDKLPKPSQKK 201

RESULT 4
RB13_HUMAN
ID RB13_HUMAN STANDARD; PRT; 203 AA.
AC P51153;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ras-related protein Rab-13.
GN RB13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94124602; PubMed=8294494;
RA Zahraoui A., Joberty G., Arpin M., Fontaine J.J., Hellio R.,
RA Tavittian A., Louvard D.;
RT "A small rab GTPase is distributed in cytoplasmic vesicles in non
RT polarized cells but colocalizes with the tight junction marker ZO-1

```

```

RT in polarized epithelial cells.";
RL J. Cell Biol. 124:101-115(1994).
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: COULD PARTICIPATE IN POLARIZED TRANSPORT, IN THE
CC ASSEMBLY AND/OR THE ACTIVITY OF TIGHT JUNCTIONS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC TIGHT JUNCTIONS OR ASSOCIATED
CC WITH VESICLES SCATTERED THROUGHOUT THE CYTOPLASM IN CELLS LACKING
CC TIGHT JUNCTIONS.
CC -1- TISSUE SPECIFICITY: VARIETY OF EPITHELIA, INCLUDING INTESTINE,
CC KIDNEY, LIVER, AND OF ENDOTHELIAL CELLS.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X75593; CAA53266.1; --
CC EMBL; BC000799; AAH00799.1; --
CC HSSP; P01112; 1PLL.
CC Genew; HGNC:9762; RAB13.
CC MIM; 602672; --
CC InterPro; IPR003579; GTPase_Rab.
CC InterPro; IPR001230; Prenyl_site.
CC InterPro; IPR001806; Ras_trnsmg.
CC InterPro; IPR005225; Small_GTP.
CC Pfam; PF00071; Ras; 1.
CC PRINTS; PR00449; RASTRNSFRMNG.
CC SMART; SM00175; RAB; 1.
CC TIGREMS; TIGR00231; small_gtp; 1.
CC GTP-binding; Lipoprotein; Prenylation; Protein transport.
CC NP_BIND 15 22 GTP (BY SIMILARITY).
CC NP_BIND 63 67 GTP (BY SIMILARITY).
CC NP_BIND 121 124 GTP (BY SIMILARITY).
CC DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
CC LIPID 200 207 GERANYL-GERANYL (BY SIMILARITY).
CC SEQUENCE 203 AA; 22774 MW; 141621CB998178DA CRC64;

Query Match 48.1%; Score 532; DB 1; Length 203;
Best Local Similarity 47.6%; Pred. No. 4.6e-36;
Matches 100; Conservative 49; Mismatches 43; Indels 18; Gaps 3;

Qy 1 MAKQYDLVFLRLIGDSGVGKTCLLCRFTDNEFHSHISTIGVDFKMTIEVDGKVIQ 60
Db 1 MAKAYDHLKLLIGDSGVGKTCLLIRFAEDFNNTYISTIGIDFKIRTVDIEGKKIKLQ 60
Qy 61 IWDTAGQERYQTITKQYRRAQIGFLVYDISERSYQHIMKWSVDVEYAPGQKILIG 120
Db 61 VMDTAGQERYQTITAYIRGAMGILYVDIIDEKSFENIQNMNKSINASAGVERLLG 120
Qy 121 NKADEQKRVGREGQOOLAKEYGMDFYETSACTNLIKESFTRLTSLVLAQHRKEGL 180
Db 121 NKCDMEKRVQKQEQADKLAREHGIRFEFSAKSSMNVDEAFSSLDIL-----LKSQ 174
Qy 181 RMRASNELALAELEEKEGKPEGPANSSKTC 210
Db 175 GRRSGN-----GNKP--PSTDLKTC 192

RESULT 5
RA01_DISOM
ID RA01_DISOM STANDARD; PRT; 200 AA.
AC P22127;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last annotation update)

```

DE Ras-related protein ORAL.
OS Discozyme ommata (Electric ray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Hypnosqualea; Pristiorajae; Batoidae;
OC Torpediniformes; Narcinoidae; Narcinidae; Discozyme.
OX NCBI_TaxID=7785;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Electric lobe;
RX MEDLINE=91115900; PubMed=1899244;
RA Ngsee J.K., Elferink L.A., Scheller R.H.;
RT "A family of ras-like GTP-binding proteins expressed in electromotor
neurons."
RL J. Biol. Chem. 266:2675-2680(1991).
CC -1- SIMILARITY: RAS-RELATED PROTEIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; M38390; AAA49230.1; -
DR PIR; A38625; A38625.
DR HSP; P05713; 3RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsfrmng.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 16 23 GTP (BY SIMILARITY).
FT NP_BIND 64 68 GTP (BY SIMILARITY).
FT NP_BIND 122 125 GTP (BY SIMILARITY).
FT DOMAIN 38 46 EFFECTOR REGION (BY SIMILARITY).
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
KW GTP-binding. 16 23 GTP (BY SIMILARITY).
FT NP_BIND 64 68 GTP (BY SIMILARITY).
FT NP_BIND 122 125 GTP (BY SIMILARITY).
FT SEQUENCE 200 AA; 22623 MW; 41D38E3D760519C5 CRC64;
Query Match 48.1%; Score 531.5; DB 1; Length 200;
Best Local Similarity 56.4%; Pred. No. 4.9e-36;
Matches 97; Conservative 44; Mismatches 30; Indels 1; Gaps 1;
QY 1 MAQY-YDVLFRLLIGDSGVGKTCVLFREFSDAFNTFTSTIGIDFKIKTVLHGKKIKL 59
Db 1 MAKTYDLLKLLIGDSGVGKTCVLFREFSDAFNTFTSTIGIDFKIKTVLHGKKIKL 60
QY 60 QIWDTAGQERYOTITKQYRRAGGIFLVYDISERSYQHIMKWSVDVEYAPGVOKILI 119
Db 61 QIWDTAGQERFHTITTSYRGAMGIMLVYDITNGKSFENISKWLRNIDEHANEDVERMLL 120
QY 120 GNKADEQKRVGREGOQOLAKKEYGMDFYETSACTNLNLIKESFTRLTSLVQL 171
Db 121 GNKCDMDKRVVYKSGEIQARHAFRTFETSAKANIINIEKAFLLAEDILQ 172
RESULT 6
RB10_CANFA STANDARD; PRT; 200 AA.
ID RB10_CANFA
AC P24409;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Ras-related protein Rab-10.
GN RAB10.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Cocker spaniel;

RX MEDLINE=91061765; PubMed=2123294;
RA Chavrier P., Vingron M., Sander C., Simons K., Zerial M.;
RT "Molecular cloning of YPT1/SEC4-related CDNs from an epithelial cell
line."
RL Mol. Cell. Biol. 10:6578-6585(1990).
CC -1- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
NEUROTRANSMITTER RELEASE.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X56387; CAA39798.1; -
DR PIR; D36364; D36364.
DR HSP; P05713; 3RAB.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsfrmng.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 16 23 GTP (BY SIMILARITY).
FT NP_BIND 64 68 GTP (BY SIMILARITY).
FT NP_BIND 122 125 GTP (BY SIMILARITY).
FT DOMAIN 38 46 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 199 199 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 200 200 GERANYL-GERANYL (BY SIMILARITY).
FT SEQUENCE 200 AA; 22569 MW; 5D52B8E8E47D4362 CRC64;
Query Match 48.0%; Score 530.5; DB 1; Length 200;
Best Local Similarity 55.8%; Pred. No. 5.9e-36;
Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;
QY 1 MAQY-YDVLFRLLIGDSGVGKTCVLFREFSDAFNTFTSTIGIDFKIKTVLHGKKIKL 59
Db 1 MAKTYDLLKLLIGDSGVGKTCVLFREFSDAFNTFTSTIGIDFKIKTVLHGKKIKL 60
QY 60 QIWDTAGQERYOTITKQYRRAGGIFLVYDISERSYQHIMKWSVDVEYAPGVOKILI 119
Db 61 QIWDTAGQERFHTITTSYRGAMGIMLVYDITNGKSFENISKWLRNIDEHANEDVERMLL 120
QY 120 GNKADEQKRVGREGOQOLAKKEYGMDFYETSACTNLNLIKESFTRLTSLVQL 171
Db 121 GNKCDMDKRVVYKSGEIQARHAFRTFETSAKANIINIEKAFLLAEDILR 172
RESULT 7
RB10_HUMAN STANDARD; PRT; 200 AA.
ID RB10_HUMAN
AC O88386; Q9D7X6;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ras-related protein Rab-10.
GN RAB10.
OS Homo sapiens (Human), and
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606, 10090;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-Human; TISSUE-Pituitary;
RX MEDLINE=20402571; PubMed=10931946;
RA Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,
Gu Y.-J., Huang C.-H., Li Y.-B., Jiang C.-L., Fu G., Zhang Q.-H.,

RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,
RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,
RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.,
RT *Gene expression profiling in the human hypothalamus-pituitary-adrenal
RT axis and full-length cDNA cloning.*;
RN Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).
[2]
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RA Wong K., Hong W., Tang B.;
RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wadatsuma M., Hosokiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yanamoto J., Wakamatsu A., Nakamura Y., Negahari K., Masuho Y.,
RA Niinomiya K., Iwayanagi T.;
RT *NEDO human cDNA sequencing project.*;
RN Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RC SPECIES=Human; TISSUE=Cervix;
RA Strausberg R.;
RN Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
[5]
RP SEQUENCE FROM N.A.
RC SPECIES=Mouse;
RA Zeng Q., Tan Y.H., Hong W.;
RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
[6]
RP SEQUENCE FROM N.A.
RC SPECIES=Mouse; STRAIN=C57BL/6J; TISSUE=Stomach;
RX MEDLINE-21085660; PubMed-11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT *Functional annotation of a full-length mouse cDNA collection.*;
RL Nature 409:685-690(2001).
CC -1- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
CC NEUROTRANSMITTER RELEASE.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF106681; AAD43034.1; .
CC EMBL; AF297660; AAG13413.1; .
CC EMBL; AK023223; BAB14474.1; .
CC EMBL; BC000896; AAH00896.1; .

DR EMBL; AF035646; AAC29313.1; .
DR EMBL; AK008725; BAB25858.1; .
DR HSSP; P05713; 3RAB.
DR Genew; HGNC:9759; RAB10.
DR MGD; MGI:105066; RAB10.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsfrmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 16 23 GTP (BY SIMILARITY).
FT NP_BIND 64 68 GTP (BY SIMILARITY).
FT NP_BIND 122 125 GTP (BY SIMILARITY).
FT DOMAIN 38 46 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 199 199 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 200 200 GERANYL-GERANYL (BY SIMILARITY).
FT CONFLICT 106 106 N -> H (IN REF. 6).
SQ SEQUENCE 200 AA; 22541 MW; 7F02B8E8E46EE1E8 CRC64;
Query Match 48.0%; Score 530.5; DB 1; Length 200;
Best Local Similarity 55.8%; Pred. No. 5.9e-36;
Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;
QY 1 MAKQ-YDVLFRLLIGDSGVGKTCCLCRFTDNEPHSHSHISTIGVDFKMTIEVDGIKVI 59
DB 1 MAKTYDLLEKLLIGDSGVGKTCVLPFRFSDDAFTFTISTIGIDFKIKIVELQKKIKL 60
QY 60 QIWDTAGQERYQITIKOYIRRAQGIIFLYVIDISSRSYQHIMKWVSVDVDEVAPEGVQKILI 119
DB 61 QIWDTAGQERFHTITTSYRGAMGIMLVYDITNGKSFENISKWLRNIDEHANEDVERMLL 120
QY 120 GNRADDEQKRVGREGQOOLAKEYGMDFYETSACTNLNKEISFRLRLVLIQ 171
DB 121 GNRCDMDKRVVPKGEQIAREHGIRFFETSAKANINIEKAFLLAEDILR 172
RESULT 8
RB8B_HUMAN STANDARD; PRT; 207 AA.
AC Q92930; Q9P293;
DT 15-JUL-1999 (Rel. 38, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ras-related protein Rab-8B.
GN RAB8B
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RA Seki N., Saito T.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 67-119 FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE-97182150; PubMed-9030196;
RA Chen D., Guo J., Gahl W.A.;
RT "RAB GTPases expressed in human melanoma cells.";
RL Biochim. Biophys. Acta 1355:1-6(1997).
CC -1- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
CC NEUROTRANSMITTER RELEASE (BY SIMILARITY).
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----


```

CC -----
DR EMBL: S53270; AAB19682.1; -
DR HSP: P05713; 3RAB.
DR MGD: MGI:96960; Mel.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001230; Prenyl_site.
DR InterPro: IPR001806; Ras_trnsfrmng.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR SMART: SM00175; RAB; 1.
DR TIGRFAMs: TIGR00231; small_GTP; 1.
KW GTP-binding; Prenylation; Lipoprotein; Protein transport;
KW Proto-oncogene.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 121 124 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 203 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 206 AA; 23557 MW; 49D832725D662942 CRC64;

Query Match 47.6%; Score 526; DB 1; Length 206;
Best Local Similarity 56.4%; Pred. No. 1.4e-35;
Matches 93; Conservative 41; Mismatches 31; Indels 0; Gaps 0;

QY 1 MAKQYDLVRLLLIGDSGVGKTCILCRFTDNEFSSHSHTIGVDPKMTIEVDGIKVRIQ 60
DB 1 MAKTYDYLKLLIGDSGVGKTCVLFPSDFSTFTIGIDPKIRTIELDGKRIKIQ 60
QY 61 IWDTAGQRYQITTKYRRAGQIFLVYDISERSYQHKMKVSDVDYAPGVGKILIG 120
DB 61 IWDTAGQREFRITTYAYRGAMGIMLYDITNEKSFNDRWIRNIEHASADVEKMIIG 120
QY 121 NKADEQKRVQREGQOQLAKEYGMDFYETSACTNLNKESFTRL 165
DB 121 NKCDVNDKRVQSKERGEKIALDYGKFMETSAKANINVENAFPL 165

RESULT 11
YPT2_SCHPO STANDARD; PRT; 200 AA.
AC P17609;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ras-related protein ypt2.
GN YPT2 OR SPAC9E9.07C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=90269232; PubMed=2112089;
RA Hengst L., Lemmelter T., Gallwitz D.;
RT "Structural and functional analysis of ypt2, an essential ras-related
RT gene in the fission yeast Schizosaccharomyces pombe encoding a Sec4
RT protein homologue.";
RL EMBO J. 9:1957-1962(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90332438; PubMed=2115995;
RA Fawell E., Hook S., Sweet D., Armstrong J.;
RT "Novel ypt1-related genes from Schizosaccharomyces pombe.";
RL Nucleic Acids Res. 18:4264-4264(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

```

```

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Welltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Shrapkovi G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC -!- TRAFFIC (BY SIMILARITY).
CC -!- MISCELLANEOUS: THIS PROTEIN IS ESSENTIAL FOR CELL VIABILITY.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC PROBABLE YEAST SEC4 HOMOLOG.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X52469; CAA36707.1; -
DR EMBL: X52864; CAA37045.1; -
DR EMBL: Z99262; CAB16405.1; -
DR PIR: S10493; S10493.
DR PIR: S12790; S12790.
DR HSP: P05713; 3RAB.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsfrmng.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR SMART: SM00175; RAB; 1.
DR TIGRFAMs: TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 16 23 GTP (BY SIMILARITY).
FT NP_BIND 64 68 GTP (BY SIMILARITY).
FT NP_BIND 122 125 GTP (BY SIMILARITY).
FT DOMAIN 38 46 EFFECTOR REGION (PROBABLE).
FT LIPID 199 199 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 200 200 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 200 AA; 22755 MW; 2C658D153A290C30 CRC64;

Query Match 46.2%; Score 510; DB 1; Length 200;
Best Local Similarity 51.9%; Pred. No. 2.6e-34;
Matches 97; Conservative 39; Mismatches 43; Indels 8; Gaps 1;

QY 3 KOYDLVRLLLIGDSGVGKTCILCRFTDNEFSSHSHTIGVDPKMTIEVDGIKVRIQ 62
DB 4 KSYDYLKLLIGDSGVGKSCILLRFSDFSTFTIGIDPKIRTIELDGKRIKIQ 63
QY 63 DTAGQERYQITTKYRRAGQIFLVYDISERSYQHKMKVSDVDYAPGVGKILGNK 122
DB 64 DTAGQERFRTITTYAYRGAMGIMLLYDITNEKSFNDRWIRNIEHASADVEKMIIG 123
QY 123 ADEQKRVQREGQOQLAKEYGMDFYETSACTNLNKESFTRLTELVLQAHKREGLRM 182

```

[illegible]

```
DR PIR: J00908; J00908.
DR PIR: J00640; J00640.
DR HSP: P05713; 3RAB.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsmg.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR SMART: SM00175; RAB; 1.
DR TIGRFAMs: TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Multigene family.
FT NP_BIND 22 29 GTP (BY SIMILARITY).
FT NP_BIND 70 74 GTP (BY SIMILARITY).
FT NP_BIND 128 131 GTP (BY SIMILARITY).
FT DOMAIN 44 52 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 213 213 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 214 214 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 216 AA; 23835 MW; 09E9C19A7A44E705 CRC64;

Query Match 46.0%; Score 508; DB 1; Length 216;
Best Local Similarity 46.9%; Pred. No. 4.2e-34;
Matches 100; Conservative 41; Mismatches 54; Indels 18; Gaps 3;

QY 5 YDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHISTIGVDFKMTIEVDGKVRQIWD 64
DB 12 YDLKLLIGDSGVGKSCLLRFSDGFTTSFTTIGIDFKIRFIELDGRKIKLIQIWD 71

QY 65 AGQERYQITKQYRRAAGIFLVYDISSERSYQHIMKWSVDVEYAPGQVOKILGNKAD 124
DB 72 AGQERFRTITAYYRGANGILLVYDVDTDESFFNIRNWRINIEHQASDNVKNILVGNKAD 131

QY 125 -EQRKQVGRGQQLAKEYGMDFYETSACTNLNKESTFTLTELVLQAHKELEGLMR 183
DB 132 MDESKRAYPTAKGQALADEYKIFETSAKTLNVEEVFSIG-----RDIQR 180

QY 184 ASNELALAE-----LEEKGKPEGPANSSKTC 210
DB 181 LSDTSRAEPATIKISQIDQAGAQATQKSC 213

RESULT 14
ID YPT2_VOLCA STANDARD; PRT; 217 AA.
AC P36861.
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE GTP-binding protein yptV2.
GN yptV2.
OS Volvox carteri.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OX Volvocaceae; Volvox.
NCBI_TaxID=3067;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=f. Nagariensis / HK10;
RX MEDLINE=94037148; PubMed=8221932;
RA Fabry S., Jacobsen A., Huber H., Palme K., Schmitt R.;
RT "Structure, expression, and phylogenetic relationships of a family of
RT ypt genes encoding small G-proteins in the green alga Volvox
RT carteri."
RL Curr. Genet. 24:229-240(1993).
CC -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC TRAFFIC (BY SIMILARITY).
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC EMBL: L08128; AAA34251.1; -.
DR PIR: S36365; S36365.
DR HSP: P05713; 3RAB.
DR InterPro: IPR003579; GTPase_Rab.
DR InterPro: IPR001806; Ras_trnsmg.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00071; ras; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR SMART: SM00175; RAB; 1.
DR TIGRFAMs: TIGR00231; small_GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport;
FT NP_BIND 20 27 GTP (BY SIMILARITY).
FT NP_BIND 68 72 GTP (BY SIMILARITY).
FT NP_BIND 126 129 GTP (BY SIMILARITY).
FT DOMAIN 42 50 EFFECTOR REGION (PROBABLE).
FT LIPID 215 215 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 216 216 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 217 AA; 24143 MW; 87D3B30C75689EAA CRC64;

Query Match 45.7%; Score 505.5; DB 1; Length 217;
Best Local Similarity 48.6%; Pred. No. 6.7e-34;
Matches 101; Conservative 40; Mismatches 64; Indels 3; Gaps 3;

QY 5 YDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHISTIGVDFKMTIEVDGKVRQIWD 64
DB 10 YDLKLLIGDSGVGKSCLLRFDDFTTSFTTIGIDFKIKVVDGKLVKLIQIWD 69

QY 65 AGQERYQITKQYRRAAGIFLVYDISSERSYQHIMKWSVDVEYAPGQVOKILGNKAD 124
DB 70 AGQERFRTITAYYRGANGILLVYDVDTDESFFNIRNWRINIEHQASDNVKNILVGNKIL 129

QY 125 -EQRKQVGRGQQLAKEYGMDFYETSACTNLNKESTFTLTELVLQAHKELEGLMR 183
DB 130 LAEDKRVVSIARGQALADEFGFRFYETSAKDNVHVEEAFIAVKDVLARMEGEHANQQL 189

QY 184 ASNELALAE-----LEEKGKPEGPANSSKTC 210
DB 190 QQQQLSAAQPVRLTSGSP-SPAQKSKC 216

RESULT 15
ID RB10_RAT STANDARD; PRT; 200 AA.
AC P35281.
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Ras-related protein Rab-10.
GN RAB10.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=92210533; PubMed=1313420;
RA Eiferink L.A., Anzal K., Scheller R.H.;
RT "Rab15, a novel low molecular weight GTP-binding protein specifically
RT expressed in rat brain."
RL J. Biol. Chem. 267:5768-5775(1992).
CC -!- FUNCTION: MAY BE INVOLVED IN VESICULAR TRAFFICKING AND
CC NEUROTRANSMITTER RELEASE.
CC -!- TISSUE SPECIFICITY: HIGHEST LEVELS IN NEURAL AND MUSCLE TISSUES.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```


GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:25:06 ; Search time 81 Seconds
(without alignments)
539.284 Million cell updates/sec

Title: US-09-817-198b-2

Perfect score: 1105

Sequence: 1 MAKQYDVLFRLLIGDSGVG.....LEEEGKPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_virus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	798	72.2	168	11 Q91YW0	Q91YW0 mus musculus
2	557	50.4	211	5 Q9TYS2	Q9TYS2 caenorhabdi
3	546.5	49.5	204	5 O15971	O15971 drosophila
4	541	49.0	207	11 Q8VCF6	Q8VCF6 mus musculus
5	528.5	47.8	207	5 O18338	O18338 drosophila
6	526.5	47.6	200	4 Q9H073	Q9H073 homo sapien
7	526	47.6	202	11 Q9DD03	Q9DD03 mus musculus
8	521	47.1	206	3 Q9HET4	Q9HET4 aspergillus
9	519	47.0	216	10 Q24466	Q24466 arabidopsis
10	518	46.9	214	10 Q40218	Q40218 lotus japon
11	515.5	46.7	216	10 Q40215	Q40215 lotus japon
12	515	46.6	216	10 Q9FJF1	Q9FJF1 arabidopsis
13	514	46.5	201	5 Q94148	Q94148 caenorhabdi
14	513.5	46.5	215	10 Q40219	Q40219 lotus japon
15	513	46.4	216	10 Q8VWF9	Q8VWF9 nicotiana t
16	512	46.3	216	10 Q40177	Q40177 lycopersico

17	512	46.3	216	10 Q9LZD4	Q9LZD4 arabidopsis
18	511.5	46.2	212	10 Q40217	Q40217 lotus japon
19	510.5	46.2	216	10 Q8W3J4	Q8W3J4 nicotiana t
20	510	46.2	218	10 Q9SF91	Q9SF91 arabidopsis
21	508.5	46.0	215	10 Q41023	Q41023 pisum sativ
22	508	46.0	216	10 Q49844	Q49844 daucus caro
23	508	46.0	216	10 Q8W3J3	Q8W3J3 nicotiana t
24	507.5	45.9	222	10 Q96362	Q96362 brassica ca
25	507	45.9	202	3 Q96VL3	Q96VL3 collettotric
26	506.5	45.8	216	10 Q9SWV8	Q9SWV8 lycopersico
27	506	45.8	216	10 Q41024	Q41024 pisum sativ
28	505	45.7	198	4 Q96GUA	Q96GUA homo sapien
29	505	45.7	216	10 Q8W3J2	Q8W3J2 nicotiana t
30	502.5	45.5	215	10 Q41022	Q41022 pisum sativ
31	495.5	44.8	204	3 Q9C1X5	Q9C1X5 pichia past
32	492.5	44.6	215	10 Q41061	Q41061 pisum sativ
33	484	43.8	205	5 Q9BLE3	Q9BLE3 entamoeba h
34	482	43.6	203	10 Q40569	Q40569 nicotiana t
35	482	43.6	205	5 Q26554	Q26554 schistosoma
36	481	43.5	203	10 Q8RU63	Q8RU63 oryza sativ
37	478.5	43.3	201	4 Q9H0U4	Q9H0U4 homo sapien
38	478	43.3	202	10 Q9ZRH6	Q9ZRH6 petunia hyb
39	477	43.2	202	3 Q9HDT5	Q9HDT5 trichoderma
40	477	43.2	202	10 Q39845	Q39845 glycine max
41	477	43.2	202	10 Q8L155	Q8L155 pisum sativ
42	476	43.1	203	10 Q940Z7	Q940Z7 arabidopsis
43	474	42.9	205	10 Q9ZRE2	Q9ZRE2 arabidopsis
44	473.5	42.9	201	11 Q9DJG1	Q9DJG1 mus musculu
45	473.5	42.9	205	5 Q9UAQ6	Q9UAQ6 caenorhabdi

ALIGNMENTS

RESULT 1

Q91YW0 ID Q91YW0 PRELIMINARY; PRT; 168 AA.
AC Q91YW0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 19.4 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC013790; AAH13790.1; -
DR InterPro; IPR001806; Ras_trnsfrimng.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR TIGRFAMs; TIGR00231; small_gtp; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
KW GTP-binding; Hypothetical protein.
SQ SEQUENCE 168 AA; 19399 MW; 4E2C2FB1C56BCDF8 CRC64;

Query Match 72.2%; Score 798; DB 11; Length 168;
Best Local Similarity 99.3%; Pred. No. 3.7e-59;
Matches 152; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHSITIGVDKFKMTIEVDGKVRQ	60
Db	1	MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHSITIGVDKFKMTIDVDGKVRQ	60
Qy	61	IWDTAGQERYQTITKQYVRRAGQIFLVYDIDISSERSYQHIMKWSVDVYAEQVOKILIG	120
Db	61	IWDTAGQERYQTITKQYVRRAGQIFLVYDIDISSERSYQHIMKWSVDVYAEQVOKILIG	120
Qy	121	NKADEQKRVQREGGQQLAKEYGMDFYTSAC	153

DR	InterPro:	IPR003579;	GTPase_Rab.
DR	InterPro:	IPR001806;	Ras_trnsmfmg.
DR	InterPro:	IPR002078;	Sig54_interact.
DR	InterPro:	IPR005225;	Small_GTP.
DR	Pfam:	PF00071;	ras; 1.
DR	PRINTS:	PR00449;	RASTRNSFRMNG.
DR	SMART:	SM00175;	RAB; 1.
DR	TIGRFAMS:	TIGR00231;	small GTP; 1.
DR	PROSITE:	PS00675;	STGMA54_INTERACT_1; UNKNOWN_1.
KW	GTP-binding;	Hypothetical protein; Lipoprotein.	
SQ	SEQUENCE	200 AA; 22469 MW; 7F01DB88E46EE3EA CRC64;	
Query Match 47.6%; Score 526.5; DB 4; Length 200;			
Best Local Similarity 55.8%; Pred. No. 2.3e-36;			
Matches 96; Conservative 45; Mismatches 30; Indels 1; Gaps 1;			
Qy	1	MAKO-YDVLFRLLLIGDSGVKTCCLCRFTDNEPHSHISTIGYDFKMKITIEVDGIKVRI	59
Dd	1	MAKTYDLLFKLLLIGDSGVKTCVLFRRSDADNTTFISITGIDFKIKTVLQGKKIKL	60
Qy	60	QIWDTAQERYQTITTKYYRQAQIGIFLYDISSERSYQHIMKWVSVDVEAPEGVQKILI	119
Dd	61	QIWDTAQGRFHPTTTTYYRGAMGIMLVYDITNGSPFNISKVLRTDEHANEDEVRLML	120
Qy	120	GKNADEPKRGVREGOQLAKEYGMDFYETSACNLNLIKESFTRLVELVLO	171
Dd	121	GKCDMDKRVVPKGSGQTAREHGIRFFETSAKANINIKAFUTLAEDILR	172
RESULT 7			
ID	Q9DD03	PRELIMINARY;	PRT; 202 AA.
QC	Q9DD03;		
DT	01-JUN-2001	(TrEMBLrel. 17, Created)	
DT	01-JUN-2001	(TrEMBLrel. 17, Last sequence update)	
DE	01-JUN-2002	(TrEMBLrel. 21, Last annotation update)	
DE	0610007N03Rlk	protein (RIKEN CDNA 0610007N03 gene).	
GN	0610007N03RK.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxId=10090;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=KIDNEY;		
RX	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana S.,		
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikolaide I., Pesole G., Quackenbush J.,		
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,		
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,		
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kotsuki S.,		
RA	Hayashizaki Y.;		
FT	"Functional annotation of a full-length mouse cdna collection."		
RL	Nature 409:685-690(2001).		
RP	[2]		
RP	SEQUENCE FROM N.A.		
RA	Strausberg R.;		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
CC	-1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.		
DR	EMBL; AK002303; BAB22000.1; -.		
DR	EMBL; BC027214; AAH27214.1; -.		

DR HSP; P05713; 3RAB.
 DR MGD; MGI.1915578; 0610007N03rik.
 DR InterPro; IPR003579; GTPase_Rab.
 DR InterPro; IPR001230; Prenyl_site.
 DR InterPro; IPR001806; Ras_trnsmg.
 DR InterPro; IPR002078; Sig54_interact.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTRNSFRMG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRFAMS; TIGR00231; small_gtp; 1.
 DR PROSITE; PS00294; PRENYLATION; UNKNOWN_1.
 DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
 KW GTP-binding; Lipoprotein.
 SQ SEQUENCE 202 AA; 22770 MW; 5DF599432E228AC0 CRC64;
 Query Match 47.6%; Score 526; DB 11; Length 202;
 Best Local Similarity 54.1%; Pred. No. 2.5e-36;
 Matches 92; Conservative 45; Mismatches 33; Indels 0; Gaps 0;
 QY 1 MAKQYDLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHISTIGVDYFKMKTIYVDGKVKRIQ 60
 DB 1 MAKAYDLFRLLIGDSGVGKTCLLIRFAEDNFNSTYISTIGIDFKIRTVDIEGKRIKQ 60
 QY 61 IWDAGQERYQITTKQYRRAQGIIFYDYDISSERSYOHIMKWYSDVDYAEVAGVQKILIG 120
 DB 61 VMDTAGQERKTIITAYIRGAGIIFYDYDISSERSYOHIMKWYSDVDYAEVAGVQKILIG 120
 QY 121 NKADEQKROVREGQOQLAKKEYGMDFYETSACTNLNLIKESFTRTLVL 170
 DB 121 NKCDMEAKROVREGQOQLAKKEYGMDFYETSACTNLNLIKESFTRTLVL 170
 RESULT 8
 Q9HET4 PRELIMINARY; PRT; 206 AA.
 AC Q9HET4;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Secretion related GTPase, (SrgA).
 OS Aspergillus niger.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5061;
 RN [1]
 RC STRAIN=N402;
 RA Punt P.J., Seiboth B., Weenink X.O., van Zeijl C.M., Lenders M.,
 RA Van Den Hondel C.A.;
 RA Konetschny C., Ram A.F., Montijn R., Kubicek C.P.,
 RA "Identification and characterisation of a family of secretion related
 RT small GTPase encoding genes from the filamentous fungus Aspergillus
 RT niger; a putative SEC4 homologue is not essential for growth.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 CC EMBL; AJ278658; CAC17832.1; -;
 DR HSP; P05713; 3RAB.
 DR InterPro; IPR003579; GTPase_Rab.
 DR InterPro; IPR001806; Ras_trnsmg.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTRNSFRMG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRFAMS; TIGR00231; small_gtp; 1.
 KW GTP-binding; Lipoprotein.
 SQ SEQUENCE 206 AA; 22823 MW; 501916B795CF8C8C CRC64;
 Query Match 47.1%; Score 521; DB 3; Length 206;
 Best Local Similarity 60.6%; Pred. No. 6.8e-36;
 Matches 97; Conservative 30; Mismatches 33; Indels 0; Gaps 0;

QY 3 KOYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHISTIGVDYFKMKTIYVDGKVKRIQIW 62
 DB 3 NYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHISTIGVDYFKMKTIYVDGKVKRIQIW 64
 QY 63 DTAGQERYQITTKQYRRAQGIIFYDYDISSERSYOHIMKWYSDVDYAEVAGVQKILIGNK 122
 DB 63 DTAGQERYQITTKQYRRAQGIIFYDYDISSERSYOHIMKWYSDVDYAEVAGVQKILIGNK 124
 QY 123 ADEQKROVREGQOQLAKKEYGMDFYETSACTNLNLIKESF 162
 DB 123 ADEQKRAVTEGQOQLAKKEYGMDFYETSACTNLNLIKESF 164
 RESULT 9
 O24466 PRELIMINARY; PRT; 216 AA.
 ID O24466;
 AC O24466;
 DT 01-JAN-1998 (Tremblrel. 05, Created)
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE ATRAB8 (GTPase ATRAB8).
 GN RAB8 OR F4P12_310.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Bischoff F., Palme K.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RC SEQUENCE FROM N.A.
 RA Bloecker H., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X., Quetier F.,
 RA Salanoubat M.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RC SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 CC EMBL; U82434; AAB65088.1; -;
 DR EMBL; AL132966; CAB67668.1; -;
 DR HSP; P05713; 3RAB.
 DR InterPro; IPR003579; GTPase_Rab.
 DR InterPro; IPR001806; Ras_trnsmg.
 DR InterPro; IPR002078; Sig54_interact.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTRNSFRMG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRFAMS; TIGR00231; small_gtp; 1.
 DR PROSITE; PS00675; SIGMA54_INTERACT_1; UNKNOWN_1.
 KW GTP-binding; Lipoprotein.
 SQ SEQUENCE 216 AA; 23939 MW; 1648E45B29D4EEB6 CRC64;
 Query Match 47.0%; Score 519; DB 10; Length 216;
 Best Local Similarity 51.9%; Pred. No. 1.1e-35;
 Matches 98; Conservative 42; Mismatches 43; Indels 6; Gaps 2;
 QY 5 YDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHSHISTIGVDYFKMKTIYVDGKVKRIQIWD 64
 DB 12 YDYLKLLIGDSGVGKTCLLCRFTDNEFHSHSHISTIGVDYFKMKTIYVDGKVKRIQIWD 71
 QY 65 AGQERYQITTKQYRRAQGIIFYDYDISSERSYOHIMKWYSDVDYAEVAGVQKILIGNKAD 124
 DB 72 AGQERYQITTKQYRRAQGIIFYDYDISSERSYOHIMKWYSDVDYAEVAGVQKILIGNKAD 131
 QY 125 -EQKROVREGQOQLAKKEYGMDFYETSACTNLNLIKESFTRTLVLQA-----HRKELE 178
 DB 132 MDESKRAVPSKQALADEYGMKFFETSAKTLNVEEVEFFSIADIKQRLADTDARAPQ 191

Search completed: June 18, 2003, 15:35:30
Job time : 83 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:32:26 ; Search time 27 Seconds
(without alignments)
231.024 Million cell updates/sec

Title: US-09-817-198b-2

Perfect score: 1105

Sequence: 1 MAKQYDVLFRLLIGDSVG.....LEEEGKPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs; 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	545	49.3	207	2	US-08-824-873-4
2	545	49.3	207	3	Sequence 4, Appl
3	527.5	47.7	205	2	US-09-198-184-4
4	527.5	47.7	205	2	Sequence 25, Appl
5	518.5	46.9	198	2	US-08-718-270A-25
6	518.5	46.9	198	2	Sequence 51, Appl
7	503.5	45.6	207	2	US-08-718-270A-51
8	503.5	45.6	207	2	Sequence 35, Appl
9	494.5	44.8	215	2	US-08-531-525-10
10	494.5	44.8	215	2	Sequence 10, Appl
11	486.5	44.2	194	2	US-08-718-270A-10
12	486.5	44.2	194	2	Sequence 34, Appl
13	478.5	43.3	201	2	US-08-916-901-3
14	478.5	43.3	201	2	Sequence 3, Appl
15	469	42.4	201	2	US-09-154-602-3
16	469	42.4	201	2	Sequence 8, Appl
17	463	41.9	202	2	US-08-531-525-8
18	463	41.9	202	2	Sequence 14, Appl
19	453.5	41.0	201	2	US-08-718-270A-14
20	453.5	41.0	201	2	Sequence 13, Appl
21	452.5	41.0	190	2	US-08-531-525-13
22	452.5	41.0	190	2	Sequence 13, Appl
23	443	40.1	218	3	US-08-824-873-3
24	443	40.1	218	2	Sequence 3, Appl
25	416	37.6	191	2	US-09-198-184-3
26	412.5	37.3	212	4	US-08-531-525-19
27	409.5	37.1	190	2	Sequence 19, Appl
					Sequence 3, Appl
					Sequence 67, Appl
					Sequence 1, Appl

28 409.5 37.1 190 3 US-09-198-184-1 Sequence 1, Appl
29 409.5 37.1 212 2 US-08-531-525-18 Sequence 18, Appl
30 409.5 37.1 212 2 US-08-718-270A-18 Sequence 18, Appl
31 409.5 37.1 214 2 US-08-531-525-52 Sequence 52, Appl
32 409.5 37.1 214 2 US-08-718-270A-52 Sequence 52, Appl
33 396.5 35.9 203 2 US-08-766-551-8 Sequence 8, Appl
34 387 35.0 210 2 US-08-531-525-16 Sequence 16, Appl
35 387 35.0 210 2 US-08-718-270A-16 Sequence 16, Appl
36 373 33.8 213 2 US-08-773-423-8 Sequence 3, Appl
37 373 33.8 217 2 US-08-773-423-3 Sequence 3, Appl
38 372 33.7 208 2 US-08-531-525-17 Sequence 17, Appl
39 372 33.7 208 2 US-08-718-270A-17 Sequence 17, Appl
40 368.5 33.3 208 2 US-08-531-525-15 Sequence 15, Appl
41 368.5 33.3 208 2 US-08-718-270A-15 Sequence 15, Appl
42 360 32.6 213 2 US-08-531-525-11 Sequence 11, Appl
43 360 32.6 213 2 US-08-718-270A-11 Sequence 11, Appl
44 359.5 32.5 191 2 US-08-531-525-26 Sequence 26, Appl
45 359.5 32.5 191 2 US-08-718-270A-26 Sequence 26, Appl

ALIGNMENTS

RESULT 1
US-08-824-873-4
; Sequence 4, Application US/08824873
; Patent No. 5843717
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Guegler, Karl
; TITLE OF INVENTION: NOVEL RAB PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,873
; FILING DATE: Filed Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0240 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 207 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 234746
US-08-824-873-4

Query Match 49.3%; Score 545; DB 2; Length 207;
Best Local Similarity 52.9%; Pred. No. 2.2e-52;
Matches 99; Conservative 45; Mismatches 43; Indels 0; Gaps 0;

RESULT. 3

: GENERAL INFORMATION:

```

;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: Patentin Release #1.0, Version #1.300
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/531,525
;
; FILING DATE: 21-SEP-1995
;
; CLASSIFICATION: 530
;

```

```

, NAME: Ferber, Donna M.
, REGISTRATION NUMBER: 33,878
, REFERENCE/DOCKET NUMBER: 37-94
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: (303) 499-8080
, TELEFAX: (303) 499-8089
, INFORMATION FOR SEQ ID NO: 25:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 205 amino acids
, TYPE: amino acid
, STRANDEDNESS: single
, TOPOLOGY: linear
, MOLECULE TYPE: protein
, HYPOTHETICAL: NO
, ORIGINAL SOURCE:
, ORGANISM: Canis familiaris
US-08-531-525-25

```

Qy	3	KQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGVDFRMKTIEVDGIKVRIQIW	62
		- - - - - - - - - - - - - - - : : : : - - - - - - - - - - - - - - -	
Db	2	KTYDYLFKLLIGSGVGKTCVLFSEDAFNSTFISIDFKIETRIELDGKRKLQIW	61

QY	3	KOYDVLFRLLIGDGVGKTCLLCRFTDNEFHS	SHISTIGVDFKMTKITEVDGKIVRQIW	62
Db	2	KYDYDLKLLIGDGVGKTVLFRFSDAENSTFISTIGD	KIRTIELDKRRIKLIW	61
QY	63	DTAGQRYQTITKQYRRAQGIWLVYDILSSRSQ	HIHKWVSVDVEYAPGVGKILGNK	122
Db	62	DTAGQRFKITIATY-RANGIMLVYDITNEKSPDN	RNRWIRIEEHASADVKEMLGNK	120
QY	123	ADEQKQYVGRGQOOLAKYGMDFYETSACTNINIK	ESFTRJUTVLQAHRELEGLRM	182

Db 121 CDVNDKQVSKERGEKALDYGIKFMETSAKANINVENAFTLARDIKAKMDKKLEGNSP 180
Qy 183 RASNE 187
Db 181 QGSNQ 185

RESULT 4

US-08-718-270A-25
; Sequence 25, Application US/08718270A
; Patent No. 5910478
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5910478le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptidomimetics Inhibiting
; TITLE OF INVENTION: the Oncogenic Action of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,270A
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
; FILING DATE: 21-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,091
; FILING DATE: 21-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8089
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 205 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Canis familiaris
US-08-718-270A-25

Query Match 47.7%; Score 527.5; DB 2; Length 205;
Best Local Similarity 52.4%; Pred. No. 1.9e-50;
Matches 97; Conservative 45; Mismatches 42; Indels 1; Gaps 1;
Qy 3 KOYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHSTIGVDFKMKTIKIEVDGKIVRIQW 62
Db 2 KYDYLKLLIGDSGVGKTCVLFREFSDAFNSTFTIGIDFKIRTIKIEVDGKIRKIQW 61
Qy 63 DTAGQERYQITTKQYRRAGQIFLVYDISSERSYOHIMKWVSDVDEYAPGEGVOKILGNK 122
Db 62 DTAGQERFRITAYY-RANGIMLVYDITNEKSPDNRNWIRNIEEHASADVEKWLGNK 120

Qy 123 ADEBQKQVREGQQQALAKYGMDFYETSACTNNIKESFTRLVQLQARRKELEGIRM 182
Db 121 CDVNDKQVSKERGEKALDYGIKFMETSAKANINVENAFTLARDIKAKMDKKLEGNSP 180
Qy 183 RASNE 187
Db 181 QGSNQ 185

RESULT 5

US-08-531-525-51
; Sequence 51, Application US/08531525
; Patent No. 5840683
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5840683le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptidomimetics Inhibiting
; TITLE OF INVENTION: the Oncogenic Action
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,525
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 37-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8089
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Canis familiaris
US-08-531-525-51

Query Match 46.9%; Score 518.5; DB 2; Length 198;
Best Local Similarity 55.0%; Pred. No. 1.8e-49;
Matches 94; Conservative 45; Mismatches 31; Indels 1; Gaps 1;

Qy 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHSHSTIGVDFKMKTIKIEVDGKIVRIQ 60
Db 1 MKKYDVLFRLLIGDSGVGKTCVLFREFSDAFNSTFTIS-IGIDFKITVLEQKKIKIQ 59
Qy 61 IWDTAGQERYQITTKQYRRAGQIFLVYDISSERSYOHIMKWVSDVDEYAPGEGVOKILIG 120
Db 60 IWDTAGQERFHTITTSYIRGAMGIMLVYDITNGKSFENISKWLNRIDEHANEDVERMLG 119
Qy 121 NKADEQKQVREGQQQALAKYGMDFYETSACTNNIKESFTRLVQLQARRKELEGIRM 171
Db 120 NKCDMDKRVVPVKGQIAHEGIRFETSAKYNINIERAFTLAEILR 170

RESULT 6

US-08-718-270A-51
; Sequence 51, Application US/08718270A
; Patent No. 5910478
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5910478le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptidomimetics Inhibiting
; the Oncogenic Action of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,270A
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
; FILING DATE: 21-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,091
; FILING DATE: 21-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Canis familiaris

US-08-718-270A-51

Query Match 46.9%; Score 518.5; DB 2; Length 198;
Best Local Similarity 55.0%; Pred. No. 1.8e-49;
Matches 94; Conservative 45; Mismatches 31; Indels 1; Gaps 1;
QY 1 MAKQYDLFRLLIGDSGVGKTCLLCRFTDNEFHSSHSITIGVDFKMTIEVDGKVRIO 60
DB 1 MKTYDLFRLLIGDSGVGKTCVLFPSDDAFNTFIS-IGIDFKITVLEGGKIKIQ 59
QY 61 IMDTAGQRYQTITKQYRRAQGIYLYVDISSERSYQHIMKWSVDVEYAPGVQKILIG 120
DB 60 INDTAQGRFHTITTSYRGAMGIMLYDITNGKSPENISKWLNRIDEHANEDVERMLIG 119
QY 121 NKADEQKRVQREGQQQLAKKEYGDMDFYETSACTNLIKESFTRLTLVLQ 171
DB 120 NKCDMDKRVKPGKEQIAREHGIRFFETSARKVININIEKAFITLAEIDILR 170

RESULT 7

RESULT 8

US-08-718-270A-35
; Sequence 35, Application US/08718270A
; Patent No. 5910478
; GENERAL INFORMATION:

US-08-531-525-35
; Sequence 35, Application US/08531525
; Patent No. 5840683
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5840683le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action
; the Oncogenic Action of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,525
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 37-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 207 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Discopyge ommata

US-08-531-525-35

Query Match 45.6%; Score 503.5; DB 2; Length 207;
Best Local Similarity 46.4%; Pred. No. 8.8e-48;
Matches 96; Conservative 56; Mismatches 44; Indels 11; Gaps 4;
QY 3 KOYDLFRLLIGDSGVGKTCLLCRFTDNEFHSSHSITIGVDFKMTIEVDGKVRIO 62
DB 2 KTYDLFRLLIGDSGVGKTCVLFPSDDAFNTFIS-IGIDFKITVLEGGKIKIQ 61
QY 63 DTAGQRYQTITKQYRRAQGIYLYVDISSERSYQHIMKWSVDVEYAPGVQKILIGNK 122
DB 62 DTAGQRYQTITKQYRRAQGIYLYVDISSERSYQHIMKWSVDVEYAPGVQKILIGNK 119
QY 123 ADEQKRVQREGQQQLAKKEYGDMDFYETSACTNLIKESFTRLTLVLQAHKEGLELRM 182
DB 120 CDMNEKRVQREGKEGLAIDYGIKLETSKSNINIEHASSDVERMILGNK 174
QY 183 RASNELALAE-LEEDEGKPEGPANSK 208
DB 175 ---NENSLQEAVDKLPKPKPKSKK 198

RESULT 8

US-08-718-270A-35
; Sequence 35, Application US/08718270A
; Patent No. 5910478
; GENERAL INFORMATION:

```

; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5910478le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptidomimetics Inhibiting
; the Oncogenic Action of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,270A
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
; FILING DATE: 21-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,091
; FILING DATE: 21-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8089
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 207 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Drosophila ommata
; US-08-718-270A-35

```

```

Query Match 45.68; Score 503.5; DB 2; Length 207;
Best Local Similarity 46.4%; Pred. No. 8.8e-46;
Matches 96; Conservative 56; Mismatches 44; Indels 11; Gaps 4;

Qy 3 KOYDLVRLLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMTIEVDGKVRQIWI 62
Db 2 KTYDLPKLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMTIEVDGKVRQIWI 61

Qy 63 DTAGQERYQITKQYRRAGIFLVYDISERSYQHIMKWSVDVYAPGVOKILGNK 122
Db 62 DTAGQERFRTIT-AYYGAMGIMKV-DITNEKSPDNKINIRNIEEHASSDVERMILGNK 119

Qy 123 ADEQKRVGREGQOQLAKKEYGMDFYETSACTNLNIKESFTRLTELVLQAHKKEGLEGRM 182
Db 120 CDMEKQVKEGKEGLAIDYGIKFLTSKSSINVEEAFITLARDINTLKNKKM----- 174

Qy 183 RASNELALAE-LEEEGKPEGPANSSK 208
Db 175 ---NENSLQEAVDKLSPPKPKSOKK 198

```

```

RESULT 9
US-08-531-525-10
; Sequence 10, Application US/08531525

```

```

; Patent No. 5840683
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5840683le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action
; the Oncogenic Action of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,525
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 37-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8089
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Arabidopsis thaliana
; US-08-531-525-10

```

```

Query Match 44.88; Score 494.5; DB 2; Length 215;
Best Local Similarity 46.9%; Pred. No. 9.2e-47;
Matches 100; Conservative 40; Mismatches 54; Indels 19; Gaps 4;

Qy 5 YDVLFRLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMTIEVDGKVRQIWI 64
Db 12 YDYLKLLIGDSGVGKTCCLCRFTDNEFHSSHSITIGVDFKMTIEVDGKVRQIWI 71

Qy 65 AGQERYQITKQYRRAGIFLVYDISERSYQHIMKWSVDVYAPGVOKILGNK 124
Db 72 AGQER-RTITAYRGAMGILLVYDVTDESFNIRNIRNIEEHASSDVERMILGNK 130

Qy 125 -EEOKRVGREGQOQLAKKEYGMDFYETSACTNLNIKESFTRLTELVLQAHKKEGLEGRM 183
Db 131 MDESKRAVPTAKQALADEYGIKFLTSKSSINVEEAFITLARDINTLKNKKM----- 179

Qy 184 ASNELALAE-----LEEEGKPEGPANSSK 210
Db 180 LSDTDSRAEPATIKISQTDQAAGAQATQKSAC 212

```

```

RESULT 10
US-08-718-270A-10
; Sequence 10, Application US/08718270A
; Patent No. 5910478
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.

```

APPLICANT: No. 59104781e, John F.
APPLICANT: Abajian, Henry B.
APPLICANT: Kende, Andrew S.
TITLE OF INVENTION: Peptidomimetics Inhibiting
the Oncogenic Action of P21 Ras
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/718,270A
FILING DATE: 20-SEP-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/531,525
FILING DATE: 21-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/004,091
FILING DATE: 21-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 78-95
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 215 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
US-08-718-270A-10

Query Match 44.8%; Score 494.5; DB 2; Length 215;
Best Local Similarity 46.9%; Pred. No. 9,2e-47;
Matches 100; Conservative 40; Mismatches 54; Indels 19; Gaps 4;
QY 5 YDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHISTIGVDFKMKTIYVDGKIKVRIQWDT 64
DB 12 YDYLKLLIGDSGVGKSCLLRFSDGSFTTSFTTIGDIFKIRTIYELDKRIKLIQWDT 71
QY 65 AQERYQITKQYRRAGIFLVYDISSERSYQHIMKWSVDVEYAPGVQKILGNAD 124
DB 72 AQGER-RATTAYTAYRGAMGILLYDVTDESSFNIRNIRIEQHASDNVKNILVGNAD 130
QY 125 -EEQKRVGREGQOQLAKYGMDFYETSACTNLNLIKESFTRTELVLQAHKRELGLRMR 183
DB 131 MDESRAVPTAKQALADEYGIKFETSAKTNLNVVEVFFSIG-----RDIQR 179
QY 184 ASNELALAE-----LEEERKEGPANSSKTC 210
DB 180 LSDTDSRAEPATIKISQTDQAAGAGQATQKSAC 212

RESULT 11
US-08-531-525-34
Sequence 34, Application US/08531525
Patent No. 5840683
GENERAL INFORMATION:

APPLICANT: Hlavka, Joseph J.
APPLICANT: Pincus, Matthew R.
APPLICANT: No. 58406831e, John F.
APPLICANT: Abajian, Henry B.
APPLICANT: Kende, Andrew S.
TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee and Winner, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/531,525
FILING DATE: 21-SEP-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 37-94
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 194 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Discopyge ommata
US-08-531-525-34

Query Match 44.2%; Score 488.5; DB 2; Length 194;
Best Local Similarity 55.8%; Pred. No. 3.6e-46;
Matches 96; Conservative 41; Mismatches 30; Indels 5; Gaps 4;
QY 1 MAKQ-YDVLFRLLIGDSGVGKTCLLCRFTDNEFHSHISTIGVDFKMKTIYVDGKIKVRI 59
DB 1 MAKYDILFRLLIGDSGVGKTCVIFRSDDAFTTISTIGDIFKINTVELGKKIKL 60
QY 60 QIMDTAGQERYQITKQYRRAGIFLVYDISSERSYQHIMKWSVDVEYAPGVQKILI 119
DB 61 QIMDTAGQERYQIT-SYRGAMGIMLVYDITNAKSFENISKLANIDEHANEDVERMLL 119
QY 120 GNKADEQRQVGRGQOQLAKYGMDFYETSACTNLNLIKESFTRTELVLQ 171
DB 120 GNK-DMEDKRVVLKSKGQ--IAEHAIRFFETSAKANINIERAKFLAEDILQ 168
RESULT 12
US-08-718-270A-34
Sequence 34, Application US/08718270A
Patent No. 5910478
GENERAL INFORMATION:
APPLICANT: Hlavka, Joseph J.
APPLICANT: Pincus, Matthew R.
APPLICANT: No. 59104781e, John F.
APPLICANT: Abajian, Henry B.
APPLICANT: Kende, Andrew S.
TITLE OF INVENTION: Peptidomimetics Inhibiting
the Oncogenic Action of P21 Ras
NUMBER OF SEQUENCES: 52


```
;
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/154,602
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/916,901
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0367 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LIVRTUT04
; CLONE: 2514506
;
US-09-154-602-3

Query Match 43.3%; Score 478.5; DB 4; Length 201;
Best Local Similarity 45.0%; Pred. No. 4.9e-45;
Matches 91; Conservative 42; Mismatches 64; Indels 5; Gaps 1;

QY 1 MAKQYDVLFRLLIGDSCVGKTCLLCRFTDNEFHSSHSITIGVDFPKMTIEVDGKVRIG 60
Db 1 MNPEYDYLKLLIGDSCVGKSCLLLRFPADDTYTESYSTIGVDFKRTIELDGKTIKLG 60
QY 61 IWDTAGQERYQITIKQYRRAGGIFLVYDIDISSYQHIMKWSVDVEYAPGVOKILIG 120
Db 61 IWDTAGQERFRTITSSYRGAGHGIIVYDVTDOESYANVKWLQEIADRYASENVNKLVG 120
QY 121 NKADEQKRVQREGQOQLAKKEYGMDFYETSACTNLNKESEFTRTLVLQAHKLEBGL 180
Db 121 NKSDLTTRKVVNDNTAKEFADSLGIPFLETSAKNATNVEQAF-----MTMAAEIKRMGP 175
QY 181 RMRASNELALAELEEKGPEG 202
Db 176 GAASGERPNLKIDSTPVKPG 197

RESULT 15
US-08-916-901-8
; Sequence 8, Application US/08916901
; Patent No. 5892012
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: RAB PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,901
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0367 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 57006
;
US-08-916-901-8

Query Match 42.4%; Score 469; DB 2; Length 201;
Best Local Similarity 51.2%; Pred. No. 5.5e-44;
Matches 83; Conservative 36; Mismatches 43; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSCVGKTCLLCRFTDNEFHSSHSITIGVDFPKMTIEVDGKVRIG 60
Db 1 MNPEYDYLKLLIGDSCVGKSCLLLRFPADDTYTESYSTIGVDFKRTIELDGKTIKLG 60
QY 61 IWDTAGQERYQITIKQYRRAGGIFLVYDIDISSYQHIMKWSVDVEYAPGVOKILIG 120
Db 61 IWDTAGQERFRTITSSYRGAGHGIIVYDVTDOESYANVKWLQEIADRYASENVNKLVG 120
QY 121 NKADEQKRVQREGQOQLAKKEYGMDFYETSACTNLNKESEF 162
Db 121 NKSDLTTRKVVNDNTAKEFADSLGVPFLETSAKNATNVEQAF 162

Search completed: June 18, 2003, 15:36:59
Job time : 37 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:35:37 ; Search time 49 Seconds
(without alignments)
468.160 Million cell updates/sec

Title: US-09-817-198b-2

Perfect score: 1105
Sequence: 1 MAKQYDLVFRLLIGDSVG.....LEEEGKPEGPANSSKTCWC 212

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1105	100.0	212	10	US-09-817-198a-2
2	1105	100.0	401	9	US-09-764-868-701
3	1092	98.8	218	10	US-09-817-198a-5
4	1077	97.5	212	10	US-09-817-198a-4
5	832	75.3	188	9	US-09-764-868-1120
6	532	48.1	246	10	US-09-925-302-534
7	530.5	48.0	218	10	US-09-925-300-1571
8	528.5	47.8	207	10	US-09-794-257-8
9	512	46.3	190	10	US-09-822-860-5
10	497	45.0	162	10	US-09-834-765-766
11	478.5	43.3	201	10	US-09-967-736-3
12	478.5	43.3	224	9	US-10-102-806-466
13	469	42.4	201	10	US-09-867-736-8
14	467	42.3	222	10	US-09-820-003A-4
15	434.5	39.3	198	10	US-09-794-257-16
16	434.5	39.3	198	10	US-09-945-173-5
17	434.5	39.3	198	10	US-09-972-529-4
18	430	38.9	223	10	US-09-817-198a-4
19	429	38.8	223	10	US-09-817-199A-2

20	427	38.6	226	9	US-09-764-868-584	Sequence 684, App
21	426	38.6	222	9	US-09-764-868-1106	Sequence 1106, Ap
22	416	37.6	191	10	US-09-794-257-14	Sequence 14, Appl
23	416	37.6	191	12	US-10-051-986-3	Sequence 3, Appli
24	412.5	37.3	212	10	US-09-350-874-67	Sequence 67, Appl
25	399	36.1	307	9	US-09-764-868-1100	Sequence 1100, Ap
26	399	36.1	312	10	US-09-925-302-783	Sequence 783, App
27	376	34.0	213	10	US-09-794-257-5	Sequence 5, Appli
28	374.5	33.9	222	9	US-09-764-868-1112	Sequence 1112, Ap
29	374.5	33.9	225	9	US-09-764-868-692	Sequence 692, App
30	374	33.8	213	9	US-10-036-542-64	Sequence 64, Appl
31	373	33.8	213	10	US-09-988-974-8	Sequence 8, Appli
32	373	33.8	217	9	US-10-036-542-89	Sequence 89, Appli
33	373	33.8	217	10	US-09-988-974-3	Sequence 3, Appli
34	373	33.8	239	10	US-09-925-301-1077	Sequence 1077, Ap
35	370	33.5	201	10	US-09-822-860-2	Sequence 2, Appli
36	353.5	32.0	624	10	US-09-834-765-5	Sequence 5, Appli
37	353.5	32.0	625	10	US-09-834-765-762	Sequence 762, App
38	353.5	32.0	832	10	US-09-834-765-2	Sequence 2, Appli
39	351	31.8	168	10	US-09-834-765-765	Sequence 765, App
40	343	31.0	161	10	US-09-834-765-763	Sequence 763, App
41	339	30.7	208	9	US-10-108-605-45	Sequence 45, Appl
42	338.5	30.6	216	10	US-09-945-173-10	Sequence 10, Appl
43	338.5	30.6	217	10	US-09-925-300-1364	Sequence 1364, Ap
44	323.5	29.3	211	12	US-10-051-986-6	Sequence 6, Appli
45	321	29.0	173	10	US-09-820-003A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-817-198A-2

; Sequence 2, Application US/09817198A

; Patent No. US20020146758A1

; GENERAL INFORMATION:

; APPLICANT: YE, Jane et al.

; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE

; FILE REFERENCE: CLO01188

; CURRENT APPLICATION NUMBER: US/09/817,198A

; CURRENT FILING DATE: 2001-03-27

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 212

; TYPE: PRT

; ORGANISM: Human

; US-09-817-198A-2

Query Match 100.0%; Score 1105; DB 10; Length 212;
Best Local Similarity 100.0%; Pred. No. 3.6e-96;
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAKQYDLVFRLLIGDSVGKTCCLCRFTDNFHHSHHISTIGVDFKMKTI	EVGDKIVRIQ	60
DB	1	MAKQYDLVFRLLIGDSVGKTCCLCRFTDNFHHSHHISTIGVDFKMKTI	EVGDKIVRIQ	60
QY	61	IWDTAGQERYOTITKQYYRRAQGIPLVYDISERSYQHIMKWSVDV	YAPGKILIG	120
DB	61	IWDTAGQERYOTITKQYYRRAQGIPLVYDISERSYQHIMKWSVDV	YAPGKILIG	120
QY	121	NKADEEQKRVGREGQQQLAKEYGMDFYETACTNLNKESTFRLTEL	VLQAHRELEGL	180
DB	121	NKADEEQKRVGREGQQQLAKEYGMDFYETACTNLNKESTFRLTEL	VLQAHRELEGL	180
QY	181	RRASNELALAELEEEGKPEGPANSSKTCWC		212
DB	181	RRASNELALAELEEEGKPEGPANSSKTCWC		212

RESULT 2

```
US-09-764-868-701
; Sequence 701, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 701
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-868-701

Query Match      100.0%; Score 1105; DB 9; Length 401;
Best Local Similarity 100.0%; Pred. No. 8e-96;
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSVGKTCCLCRFTDNEFHSHSHISTIGVDFPKMTIEVDGKVRQ 60
DB 27 MAKQYDVLFRLLIGDSVGKTCCLCRFTDNEFHSHSHISTIGVDFPKMTIEVDGKVRQ 86
QY 61 IWDTAGQERYQTITKQYRRAGGIFLVYDISERSYQHIMKWSDVDEYAPGVQKILIG 120
DB 87 IWDTAGQERYQTITKQYRRAGGIFLVYDISERSYQHIMKWSDVDEYAPGVQKILIG 146
QY 121 NKADEEQKRVGREGQOQLAKEYGMDFYETSACTNLNIKESFTRTLVLQAHRLDGL 180
DB 147 NKADEEQKRVGREGQOQLAKEYGMDFYETSACTNLNIKESFTRTLVLQAHRLDGL 206
QY 181 RMRASNELALAELEEEGKPEGPANSSKTCWC 212
DB 207 RMRASNELALAELEEEGKPEGPANSSKTCWC 238

RESULT 3
US-09-817-198A-5
; Sequence 5, Application US/09817198A
; Patent No. US20020146758A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001188
; CURRENT APPLICATION NUMBER: US/09/817,198A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Human
US-09-817-198A-5

Query Match      98.8%; Score 1092; DB 10; Length 218;
Best Local Similarity 97.2%; Pred. No. 6.2e-95;
Matches 212; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

QY 1 MAKQYDVLFRLLIGDSVGKTCCLCRFTDNEFHSHSHISTIGVDFPKMTIEVDGKVRQ 60
DB 1 MAKQYDVLFRLLIGDSVGKTCCLCRFTDNEFHSHSHISTIGVDFPKMTIEVDGKVRQ 60
QY 61 IWDTAGQERYQTITKQYRRAGGIFLVYDISERSYQHIMKWSDVDEYAPGVQKILIG 120
DB 61 IWDTAGQERYQTITKQYRRAGGIFLVYDISERSYQHIMKWSDVDEYAPGVQKILIG 120
QY 121 NKADEEQKRVGREGQOQLAKEYGMDFYETSACTNLNIKESFTRTLVLQAHRLDGL 174
DB 121 NKADEEQKRVGREGQOQLAKEYGMDFYETSACTNLNIKESFTRTLVLQAHRLDGL 174

US-09-764-868-1120
; Sequence 1120, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1120
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (139)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (151)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE

US-09-817-198A-4
; Sequence 4, Application US/09817198A
; Patent No. US20020146758A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001188
; CURRENT APPLICATION NUMBER: US/09/817,198A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-817-198A-4

Query Match      97.5%; Score 1077; DB 10; Length 212;
Best Local Similarity 97.6%; Pred. No. 1.5e-93;
Matches 207; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAKQYDVLFRLLIGDSVGKTCCLCRFTDNEFHSHSHISTIGVDFPKMTIEVDGKVRQ 60
DB 1 MAKQYDVLFRLLIGDSVGKTCCLCRFTDNEFHSHSHISTIGVDFPKMTIEVDGKVRQ 60
QY 61 IWDTAGQERYQTITKQYRRAGGIFLVYDISERSYQHIMKWSDVDEYAPGVQKILIG 120
DB 61 IWDTAGQERYQTITKQYRRAGGIFLVYDISERSYQHIMKWSDVDEYAPGVQKILIG 120
QY 121 NKADEEQKRVGREGQOQLAKEYGMDFYETSACTNLNIKESFTRTLVLQAHRLDGL 180
DB 121 NKADEEQKRVGREGQOQLAKEYGMDFYETSACTNLNIKESFTRTLVLQAHRLDGL 180
QY 181 RMRASNELALAELEEEGKPEGPANSSKTCWC 212
DB 181 RMRASNELALAELEEEGKPEGPANSSKTCWC 212

RESULT 5
US-09-764-868-1120
; Sequence 1120, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1120
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (139)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (151)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
```

; LOCATION: (161)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (164)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (188)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; US-09-764-868-1120

Query Match 75.3%; Score 832; DB 9; Length 188;
Best Local Similarity 97.6%; Pred. No. 1.4e-70;
Matches 160; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 MAKQYDLVRLLLIGDSGVGKTCCLCRFTDNEFHSHISTIGVDFKMKTIYVDGKIVRIQ 60
Db 24 MAKQYDLVRLLLIGDSGVGKTCCLCRFTDNEFHSHISTIGVDFKMKTIYVDGKIVRIQ 83
QY 61 INDTAGQERYQTITKQYRRAGIFLVYDISSERSYOHIMKWVSDVDEYAPGVQKILIG 120
Db 84 INDTAGQERYQTITKQYRRAGIFLVYDISSERSYOHIMKWVSDVDEYAPGVQKILIG 143
QY 121 NKADEEQKRVGREGOQOLAKEYGMDFYETSACTNLNIKESFTR 164
Db 144 NKADEEQKRVGREGOQOLAKEYGMDFYETSACTNLNIKESFTR 187

RESULT 6

US-09-925-302-534
; Sequence 534, Application US/09925302
; Patent No. US20020044941A1

GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA104

; CURRENT APPLICATION NUMBER: US/09/925,302

; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05918

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 896

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 534

; LENGTH: 246

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-925-302-534

Query Match 48.1%; Score 532; DB 10; Length 246;
Best Local Similarity 47.6%; Pred. No. 3e-42;
Matches 100; Conservative 49; Mismatches 43; Indels 18; Gaps 3;

QY 1 MAKQYDLVRLLLIGDSGVGKTCCLCRFTDNEFHSHISTIGVDFKMKTIYVDGKIVRIQ 60
Db 44 MAKAYDHLFLLLLIGDSGVGKTCCLIRFADFNPNYISTIGIDFKIRIVDGGKIKIQ 103
QY 61 INDTAGQERYQTITKQYRRAGIFLVYDISSERSYOHIMKWVSDVDEYAPGVQKILIG 120
Db 104 VMDTAGQERFKTITTAIRGAMGILVYDITDEKSPENIQNMWKSINASAGVERLLIG 163
QY 121 NKADEEQKRVGREGOQOLAKEYGMDFYETSACTNLNIKESFTRLTVELVLAHRKEGL 180
Db 164 NKCDMEARKVQKEQADKLAREHGIRFFETSAKSSMNVDEAFSSLRDIL-----LKSG 217
QY 181 RNASNELALAELEEKGKPGPANSKTC 210
Db 218 GRRSGN-----GNKP--PSTDLCYC 235

RESULT 7

US-09-925-300-1571
; Sequence 1571, Application US/09925300

; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1571
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-925-300-1571

Query Match 48.0%; Score 530.5; DB 10; Length 218;
Best Local Similarity 55.8%; Pred. No. 3.6e-42;
Matches 96; Conservative 46; Mismatches 29; Indels 1; Gaps 1;
QY 1 MAKQ-YDVLFRLLLLIGDSGVGKTCCLCRFTDNEFHSHISTIGVDFKMKTIYVDGKIVRI 59
Db 19 MAKKTYDLLFKLLLLIGDSGVGKTCVLFREFSDAFTTISTIGIDFKIKTVLQGGKKIKL 78
QY 60 QIWDTAGQERYQTITKQYRRAGIFLVYDISSERSYOHIMKWVSDVDEYAPGVQKILI 119
Db 79 QIWDTAGQERFHTTTSYRGAMGIMLVYDITGKSFENISKWLRNIDEHANEDVERMLL 138
QY 120 GNKADEEQKRVGREGOQOLAKEYGMDFYETSACTNLNIKESFTRLTVELVQ 171
Db 139 GNKCDMDKRVVPKGEQIAREHGIRFFETSAKANINIEKAFITLAEDILR 190

RESULT 8

US-09-794-257-8
; Sequence 8, Application US/09794257
; Patent No. US2002009804A1

GENERAL INFORMATION:

; APPLICANT: Meyers, Rachel

; TITLE OF INVENTION: 32705, 23224, 27423, 32700, 32712, No. US2002009804A1el

; FILE REFERENCE: Human G-Proteins

; FILE REFERENCE: 35800/209285

; CURRENT APPLICATION NUMBER: US/09/794,257

; CURRENT FILING DATE: 2001-02-27

; PRIOR APPLICATION NUMBER: 60/185,606

; PRIOR FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 207

; TYPE: PRT

; ORGANISM: homo sapiens

; US-09-794-257-8

Query Match 47.8%; Score 528.5; DB 10; Length 207;
Best Local Similarity 54.0%; Pred. No. 5.2e-42;
Matches 95; Conservative 47; Mismatches 33; Indels 1; Gaps 1;
QY 1 MAKQYDLVRLLLIGDSGVGKTCCLCRFTDNEFHSHISTIGVDFKMKTIYVDGKIVRIQ 60
Db 1 MAKTYDLFLLLLIGDSGVGKTCCLFRFSEDAFTTISTIGIDFKIRITELDGKKIKLQ 60
QY 61 INDTAGQERYQTITKQYRRAGIFLVYDISSERSYOHIMKWVSDVDEYAPGVQKILIG 120
Db 61 INDTAGQERFRTITTAIRGAMGIMLVYDITNEKSFENIKWIRNIEEHASSDVERMILG 120
QY 121 NKADEEQKRVGREGOQOLAKEYGMDFYETSACTNLNIKES-FTRLTVELVLAHRK 175
Db 121 NKCDMDKRVSKERGEKLAIDYGIKIFLETSAKSSANVEAFITLARDIMTKLNKR 176

```
RESULT 9
US-09-822-860-5
; Sequence 5, Application US/09822860
; Patent No. US20020146795A1
; GENERAL INFORMATION:
; APPLICANT: ZHU, Shaoping et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; FILE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001214
; CURRENT APPLICATION NUMBER: US/09/822,860
; CURRENT FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Discopyge ommata
US-09-822-860-5

Query Match          46.3%; Score 512; DB 10; Length 190;
Best Local Similarity 47.7%; Pred. No. 1.6e-40;
Matches 93; Conservative 52; Mismatches 42; Indels 8; Gaps 1;

QY 6 DVLFRLLIGDSGVGKTCCLCRFTDNEFHSHTIGVDFKMKTIYVDGKIKVRIQIWDTA 65
DB 1 DYLFKLLIGDSGVGKTCCLFRFSEDATFSTIGIDFKIRTVELDGKKIKIWDTA 60

QY 66 GOERYQITTKQYRRAGQIFLVYDISSERSYQHIMKWSDVDEYAPGQVQKILGNRADE 125
DB 61 GOERFRTITAYYRGAMGIMLVYDITNEKSPDNIRNIRNIEEHASDVVERMILGNKDM 120

QY 126 EQKRGVQGOQOLAKGYDMDFYETSACTNINIKESFRLVLOAHRKELEGLRMRAS 185
DB 121 NEKQVSKERGEKLAIDYGIKLETSKASSINVEAFITLARDINTKLKKM----- 172

QY 186 NELALAELEEEGKP 200
DB 173 NENSLQEAVDKLSP 187

RESULT 10
US-09-834-765-766
; Sequence 766, Application US/09834765
; Patent No. US20020055478A1
; GENERAL INFORMATION:
; APPLICANT: Mary Faris
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Arthur B. Raitano
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: GTP-BINDING PROTEIN USEFUL IN TREATMENT
; TITLE OF INVENTION: AND DETECTION OF CANCER
; FILE REFERENCE: 129.6USUI
; CURRENT APPLICATION NUMBER: US/09/834,765
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/197,647
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 770
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 766
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-765-766

Query Match          45.0%; Score 497; DB 10; Length 162;
Best Local Similarity 55.4%; Pred. No. 3.4e-39;
Matches 87; Conservative 41; Mismatches 29; Indels 0; Gaps 0;
```

```
QY 9 FRLLIGDSGVGKTCCLCRFTDNEFHSHTIGVDFKMKTIYVDGKIKVRIQIWDTAQGE 68
DB 1 FKLLIGDSGVGKTCVLFRFSEDATFSTIGIDFKIRTVELDGKKRIKLIQIWDTAQGE 60

QY 69 RYQITTKQYRRAGQIFLVYDISSERSYQHIMKWSDVDEYAPGQVQKILGNRADEQK 128
DB 61 RFTTITAYYRGAMGIMLVYDITNEKSPDNIRNIRNIEEHASDVVERMILGNKCDVNDK 120

QY 129 RQVREGOQOLAKGYDMDFYETSACTNINIKESFTRL 165
DB 121 RQVSKERGEKLAIDYGIKLETSKASSINVENAFETL 157

RESULT 11
US-09-967-736-3
; Sequence 3, Application US/09967736
; Patent No. US20020103340A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: RAB PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/967,736
; FILING DATE: 28-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/154,602
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0367 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LIVRUT04
; CLONE: 2514506
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-967-736-3

Query Match          43.3%; Score 478.5; DB 10; Length 201;
Best Local Similarity 45.0%; Pred. No. 2.5e-37;
Matches 91; Conservative 42; Mismatches 64; Indels 5; Gaps 1;
```

```
QY 1 MAKQYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHTIGVDFKMKTIYVDGKIKVRIQ 60
DB 1 MNPYDVLFRLLIGDSGVGKTCCLCRFTDNEFHSHTIGVDFKMKTIYVDGKIKVRIQ 60

QY 61 IWDTAGQERYQITTKQYRRAGQIFLVYDISSERSYQHIMKWSDVDEYAPGQVQKILIG 120
DB 61 IWDTAGQERYQITTKQYRRAGQIFLVYDISSERSYQHIMKWSDVDEYAPGQVQKILIG 120
```

Qy	121	NKADEOKRQVREGGQOLAKKEYGMDFTYETSACTNLNIKESFTTRTELVLQHRRELEGL	180
Db	121	NKSDLTTRKVVNDNTAKEFADSLGIPFLETSAKNATNVEQAF-----MTMAAEIKRMRGP	175
Qy	181	RMRASNELALAELEEEEGKPRG	202
Db	176	GAASGGGERPNLKIDSTPVKPKAG	197

RESULT 12
US-10-102-806-466
: Sequence 466, Application US/10102806
: Publication No. US20030054421A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

```

, FILE REFERENCE: PA103P1C1
, CURRENT APPLICATION NUMBER: US/10/102,806
, CURRENT FILING DATE: 2002-03-22
, PRIOR APPLICATION NUMBER: 09/925,298
, PRIOR FILING DATE: 2001-08-10
, PRIOR APPLICATION NUMBER: PCT/US00/05881
, PRIOR FILING DATE: 2000-03-08
, PRIOR APPLICATION NUMBER: 60/124,270
, PRIOR FILING DATE: 1999-03-12
, NUMBER OF SEQ ID NOS: 846
, SOFTWARE: PatentIn Ver. 2.0
, SEQ ID NO 466
, LENGTH: 224
, TYPE: PRT
, ORGANISM: Homo sapiens
US-10-102-806-466

```

Query Match 43.3%; Score 478.5; DB 9; Length 224;
Best Local Similarity 45.0%; Pred. NO. 2.9e-37;
Matches 91; Conservative 42; Mismatches 64; Indels 5

Qy	1	MAKOYDVLFRLLLLIGDSGVOKTCLLCRFTNEFHSSHISTIGVDFWKMTIEVDGKIVRIQ	60
Db	24	MNPEYDFLLKLLIGDSGVOKSCILLRFADDTVYESITIGVDFKIRTIELDGKTIKLQ	83
Qy	61	IWDPAQERQCTITIKOYRRAQGLFYVDTSSESYOHMKKWSVDVEYAPGVOKTILG	120
Db	84	IWDPAQERFRTITSSYIRGAHGLIVYDVTQDSYANVKQWQEIQDIYASENVKLLVG	143
Qy	121	NKADEEQKVGREGQOGLAKEYGMDFYETESACTNLNIKESFTLTELVLQAHKKELEGL	180
Db	144	NKSLDITTKVVDNTITAKEFADSLGIPLEISAKNATNVEQAF-----MTMAAEIKKRMP	198
Qy	181	RMRASNELALAELEEEGKPEG	202
Db	199	GAASGGRRPNLKIDSTPVKPKG	220

```

1  RESULT 13
2  US-09-967-736-8
3  : Sequence 8, Application US/09967736
4  : Patent NO. US20020103340A1
5  : GENERAL INFORMATION:
6  : APPLICANT: Hillman, Jennifer L.
7  : Lai, Pretti
8  : Corley, Neil C.
9  : Shah, Purvi
10 : TITLE OF INVENTION: RAB PROTEINS
11 : NUMBER OF SEQUENCES: 9
12 : CORRESPONDENCE ADDRESS:
13 :

```

COMPANY: INCYTE PHARMACEUTICALS, INC.
 ADDRESS: INCYTE PHARMACEUTICALS, INC.
 STREET: 3174 PORTER DR.
 CITY: PALO ALTO
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:

```

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Fast-Seq for Windows Ver
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/967,736
FILING DATE: 28-Sep-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/154,602
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0367
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0355
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 201 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 57006
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-967-736-8

```

Query Match 42.4%; Score 469; DB 10; Length 201;
Best Local Similarity 51.2%; Pred. No. 1.9e-36;
Matches 83: Conservative 36; Mismatches 43; Indels

[illegible]

RESULT 14

```

US-09-820-003A-4
; Sequence 4, Application US/098200003A
; Patent No. US20020142382A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al.
; TITLE OF INVENTION: ISOLATED HUMAN M
; TITLE OF INVENTION: NUCLEIC ACID M
; TITLE OF INVENTION: PROTEINS, AND
; FILE REFERENCE: CLO01196
; CURRENT APPLICATION NUMBER: US/09/8
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Versi
; SEQ ID NO 4
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-820-003A-4

```

Query Match 42.3%; Score 467; DB 10; Length 222;
Best Local Similarity 51.9%; Pred. No. 3.4e-36;
Matches 84; Conservative 34; Mismatches 44; Indels

Qy	1	MAKQYDVLFRLLIGDSGVGKTCLLCRFTDNEFHSSHISTIGYDFKMKITIEVDGKVRIQ	60
		: : : : : : : :	
Db	21	MNPEYDLFKLLIGDSGVGKSCLLRFADDTYTESYISTIGYDFKIRITELDGTTIKIO	80

QY 61 IWDTAQERYQITKQYRRAQGIPLVYDISSERSYQHIMKWSDVDEYAPGQKILIG 120
Db 81 IWDTAQERFRITSSYRGAHIIIVYDVTQESFNQKWLQELDRYASENVNKLIVG 140
QY 121 NKADDEQKQVREGQQLAKEYGMDFYETSACTNLNIKESF 162
Db 141 NKCDLTTRKKVDYTTAKEFADSLGIPFLETSAKNATNVEQSF 182

RESULT 15

US-09-794-257-16
; Sequence 16, Application US/09794257
; Patent No. US20020009804A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 32705, 27423, 27424, 27423, 32700, 32712, No. US20020009804A1e1
; FILE REFERENCE: 35800/209285
; CURRENT APPLICATION NUMBER: US/09/794,257
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/185,606
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam accession number PF00071
US-09-794-257-16

Query Match 39.3%; Score 434.5; DB 10; Length 198;
Best Local Similarity 50.9%; Pred. No. 3.3e-33;
Matches 89; Conservative 30; Mismatches 43; Indels 13; Gaps 3;
QY 10 RLLIGDSGVGKTCILCRFTDNEFHSSHSITIGVDFKMTIEVDGKIVRIQIWDTAGQER 69
Db 1 KLVLLIGDSGVGKSSLLIRFTDNKFVEEYIPTIGVDFTYTKTVEVDGKTVKLQIWDTAGQER 60
QY 70 YQITKQYRRAQGIPLVYDISSERSYQHIMKWSDVDEYA--PEGVOKILIGNKAD--- 124
Db 61 FRALRPAYRGAQGLLYVDITSRDSFENVKKWLEILLRHADKDENVPIVLVGNKCDLED 120
QY 125 -----EQKQVREGQQLAKEYG-MDFYETSACTNLNIKESFTRITELVLQ 171
Db 121 DEDELTGQKRVVSTEEGALAKELGALPFMETS AKTNTNVEEAFEELAREILK 175

Search completed: June 18, 2003, 15:44:58
Job time : 50 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 23, 2003, 10:45:22 ; Search time 2485 Seconds
(without alignments)
2482.815 Million cell updates/sec

Title: US-09-817-198B-2

Perfect score: 1105

Sequence: 1 MAQYDVLFLLLIGDSVG.....LEEECKPEGPANSSKTCWC 212

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q/cgn2_1/USPTO.spool/US09817198/runat_18062003_144633_27909/app_query.fasta_1.391
-DB=GenEmbl -OFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCI=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09817198 -CGN_1_1_2496_@runat_18062003_144633_27909 -NCPU=3
-NO_MMAL -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPOP=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_em.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*

RESULT 1

- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pla.*
- 35: em_htg_rpd.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1105	100.0	1054	6	AX399903 Sequence
2	1090	98.6	3139	10	BC027769 Mus muscu
3	1077	97.5	945	10	M83679 Sprague-Daw
4	798	72.2	2560	10	BC013790 Mus muscu
5	546.5	49.5	1613	3	AB006189 Drosophila
6	546.5	49.5	2349	3	AY060425 Drosophila
7	545	49.3	624	9	AF498943 Homo sapi
8	545	49.3	660	9	AF498943 Homo sapi
9	545	49.3	760	4	AF498943 Homo sapi
10	545	49.3	2048	9	BC002977 Homo sapi
11	545	49.3	2818	9	AK025165 Homo sapi
12	541	49.0	1337	10	BC019990 Mus muscu
13	539.5	48.8	765	5	DXGORA2 Discoptye o
14	534.5	48.4	1211	9	BC000799 Homo sapi
15	534.5	48.4	1238	9	HSRAB13 H.sapiens m
16	532	48.1	612	9	AF498948 Homo sapi
17	531.5	48.1	840	5	DYGORA1 Discoptye o
18	530.5	48.0	603	9	AF297660 Homo sapi
19	530.5	48.0	603	9	AF498945 Homo sapi
20	530.5	48.0	888	4	CFRAB10 Canine rabl
21	530.5	48.0	897	10	AF035646 Mus muscu
22	530.5	48.0	1029	9	BC000896 Homo sapi
23	530.5	48.0	3164	9	AF106681 Homo sapi
24	530.5	48.0	3533	9	AK023223 Homo sapi
25	529.5	47.9	954	8	ATHARA3 Arabidopsi
26	529.5	47.9	969	8	AY042795 Arabidopsi
27	529.5	47.9	1143	8	AY035132 Arabidops
28	528.5	47.8	624	6	AX236078 Sequence
29	528.5	47.8	740	10	U53475 Rattus norv
30	528.5	47.8	759	10	AF525280 Rattus no
31	528.5	47.8	1128	9	BC020654 Homo sapi
32	528.5	47.8	1161	6	AX236076 Sequence
33	528.5	47.8	1265	9	AB038995 Homo sapi
34	528.5	47.8	2048	3	AY069671 Drosophila
35	528.5	47.8	2210	3	D84347 Drosophila
36	528.5	47.8	2497	6	AX285074 Sequence
37	528.5	47.8	2497	6	AX285080 Sequence
38	528.5	47.8	2497	6	AX285089 Sequence
39	528.5	47.8	2497	6	AX285091 Sequence
40	528.5	47.8	3077	9	AK001111 Homo sapi
41	528.5	47.8	3110	9	AL833365 Homo sapi
42	526.5	47.6	1113	8	LJRA88E L.japonicus
43	526.5	47.6	3521	9	HSM801620 Homo sapi
44	526	47.6	638	10	S53270 MEL-RAS-rel
45	526	47.6	1118	10	BC027214 Mus muscu

ALIGNMENTS

```

Clone-distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 55 Row: d Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.
Location/Qualifiers
1. 3139
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:38375 IMAGE:5345297"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI_CGAP_Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
208...846
/codon_start=1
/product="Unknown (protein for MGC:38375)"
/protein_id="AAH27769.1"
/db_xref="GI:2038072"

```


Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 660)
Zahraoui,A.
Direct Submission
Submitted (26-NOV-1990) A. Zahraoui, INSERM-U 248, 10 AVENUE DE
VERDUN, 750-10 PARIS, FRANCE

2 (bases 1 to 660)
Zahraoui,A., Joberty,G., Arpin,M., Fontaine,J.J., Hellio,R.,
Tavtlian,A. and Louvard,D.
A small Rab GTPase is distributed in cytoplasmic vesicles in non
polarized cells but colocalizes with the tight junction marker ZO-1
in polarized epithelial cells
J. Cell Biol. 124 (1-2), 101-115 (1994)

J4124602
MEDLINE
PUBMED
FEATURES

Location/Qualifiers
1..660
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human pheochromocytoma cDNA library"
11..634
/gene="rab8"
11..634
/gene="rab8"
/codon_start=1
/product="rab8 small GTP binding protein"
/protein_id="CAA0065.1"
/db_xref="GI:452318"
/translation="MAKTYDYLFKKLLIGDSGVGKTCVLFPSDAFNTFTSTIGID
FKIRTELDGKRRIQLQWDTAGQPERFTITAYRGAMGLIVYDITNKSFDNRINR
INRIEPAHSADQVKMTLGNCKVDNRQVSKERGEKLALDYGIKFMTSAKANINVEN
AFPTLRDIKAKMDKKLEGNPSQGSNQGVKITPDQOKRSFFRCVLL"

BASE COUNT 197 a 159 c 172 g 132 t
ORIGIN

Alignment Scores:
Pred. No.: 3,64e-45 Length: 660
Score: 545.00 Matches: 99
Percent Similarity: 77.01% Conservative: 45
Best Local Similarity: 52.94% Mismatches: 43
Query Match: 49.32% Indels: 0
DB: 9 Gaps: 0

US-09-817-198B-2 (1-212) x HSMRAB8 (1-660)

QY 1 MetaAlaIysGlnTyrrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValGly 20
Db 11 ATGGCGAAGACCTTAAATTTAGGACCATTAGAGCTCGATCGGGGACTCGGGGGTGGG 70
QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisIleSerThr 40
Db 71 AAGACCTGTGCTCTGTTCGGCTTCGCGAGGAGCCCTCAACTCCACTTTATCTCCACC 130
QY 41 IleGlyValaspPheIysMetLysThrIleGluValaspIleLysValArgIleGln 60
Db 131 ATAGGAATTGACCTTTAAAAATTAGGACCATTAGAGCTCGATGGCAAGAAGATAAATGCAG 190
QY 61 IleTrpAspThrAlaGlyGlnGluArgTyrrGlnThrIleThrIysGlnTyrrTyrrArgArg 80
Db 191 ATATGGGACACAGCCGGTTCAGAACGGTTTCGGACCATCACACAGCCCTACTACAGGGGT 250
QY 81 AlaGlnGlyIlePheLeuValTyrrAspIleSerSerGluArgSerTyrrGlnHisIleMet 100
Db 251 GCATGGGATCATGCTGTCTACGACATCAACCAAGAGAGTCTTCGACACATCCGG 310
QY 101 LysTrpValserAspValaspGlnTyrrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db 311 AACCTGGATTGCGAACATTCAGGAGCAGCGCTCTGTCAGACGCTCGAAAAGATGATCTCGGG 370
QY 121 AsnLysAlaAspGluGlnGlnLysArgGlnValGlyArgGluGlnGlnLeuAla 140
Db 371 AACACAGTGTGTGAATGACACAGACACAAGTTTTCCAAGGACGCGGGAAGACGTGGCC 430

Qy 181 ArgMetargAlaSerAsnGlu 187
Db 617 AGCCCCAGGGGAGCAACAG 637

RESULT 11
AK025165.
LOCUS Homo sapiens cDNA: FLJ21512 fis, clone COL05769. PRI 29-SEP-2000
DEFINITION AK025165
VERSION Homo sapiens colon cDNA to mRNA, clone COL05769.
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2818)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel.81-3-5449-5286,
Fax:81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

FEATURES
source Location/Qualifiers
1 .2818
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL05769"
/tissue_type="colon"
/clone_lib="COL"
/note="cloning vector pME18SFL3"
BASE COUNT 701 a 800 c 708 g 609 t
ORIGIN
Alignment Scores:
Pred. No.: 2.15e-44 Length: 2818
Score: 545.00 Matches: 99
Percent Similarity: 77.01% Conservative: 45
Best Local Similarity: 52.94% Mismatches: 43
Query Match: 49.32% Indels: 0
DB: 9 Gaps: 0

US-09-817-198B-2 (1-212) x AK025165 (1-2818)

Qy 1 MetaLysGlnTyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyValGly 20
Db 246 ATGGCGAAGACCTACGATTACCTGTTCAAGCTGCTGCTGATCGGGAGCTCGGGGTGGG 305

Qy 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40
Db 306 AAGACCTGTGCTGCTTCGGCTTCCGAGACGCCCTCACTCCACTTTTATCTCCACC 365

Qy 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
Db 366 ATAGGAATTGACTTTAAATTTAGGACCATAGAGCTCGATGTCGAAGAAATTAACATGCAG 425

Qy 61 IleTrpAspThrAlaGlnGluArgTyrGlnThrIleThrLysGlnTyrArgArg 80
|||||

Db 426 ATATGGGACACAGCCGGTCAGGAACGGTTTCGGAGCATCAACAGCGCCTACTACAGGGGT 485

Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
|||||

Db 486 GCAATGGGCATCATGCTGCTACGACATCAACCAAGAGAGTCTCTTCGACCAACATCGG 545

Qy 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
|||||

Db 546 AACGTGGATTCCCAACATTTGAGGAGCAGCCCTCTCGACAGCTCGAAAAGATGATATCGG 605

Qy 121 AsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
|||||

Db 606 AACAAAGTGTGATGTGAATGACAGACAGCAAGTTTCCRAGGAGACGGGAGAAAGCTGCC 665

Qy 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
|||||

Db 666 CTCGACTATGGAATCAAGTTTCATGGACACCGAGCGCAAGGCAACATCAATGTGAAAAAT 725

Qy 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180
|||||

Db 726 GCATTTTTCACCTCTCGCCAGAGATCAACAAACAAAATGGACAAAATTTGGAAGGCAAC 785

Qy 181 ArgMetargAlaSerAsnGlu 187

Db 786 AGCCCCAGGGGAGCAACAG 806

RESULT 12
LOCUS BC019990
DEFINITION Mus musculus, Similar to mel transforming oncogene (derived from cell line NK14) - RAB8 homolog, clone MGC:28462 IMAGE:416110, mRNA, complete cds.
ACCESSION BC019990
VERSION BC019990.1 GI:18043408
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1337)
Strausberg,R.
Direct Submission
Submitted (19-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunnaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAP Plate: 37 Row: c Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
Location/Qualifiers
1 .1337
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:28462 IMAGE:416110"

FEATURES
source

QY 101 LysTrpValSerAspValaspGluTyAlaProGluGlyValGlnLysIleLeuIleGly 120
 Db 314 AACTGGATAAGGAATATTGAAGACATGCTTCAGATGTAGAAAGAAATCATTTCTGGC 373
 QY 121 AsnLysAlaaspGluGluGlnLysArgGlnValArgGlnGlnGlnGlnLeuAla 140
 Db 374 ACAAGTGCACATGATGAATAAGGCAAGTGTCAAGAGACGAGGAAAGATTAGCG 433
 QY 141 LysGluTyGlyMetAspPheTyGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
 Db 434 ATAGATTATGGATAAGTTCTTGGAAACCCAGCGCAAAATCCAGCATATAACGTTGAAGAG 493
 QY 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeu 180
 Db 494 GCATTTATTACATTCGCAGAGATATCATGACGAAATCAACAGAAATG----- 544
 QY 181 ArgMetArgAlaSerAsnGluLeuAlaLeuAlaGlu---LeuGluGluGluGluGlyLys 199
 Db 545 -----AATGAACAGTCTCCAGGAGCAGTGCACAAAGTTAAATACCG 589
 QY 200 ProGluGlyProAlaAsnSerSerLys 208
 Db 590 CCAAGAAACCCAGTCAAAAGAAAAG 616

RESULT 14

BC000799 1211 bp mRNA linear PRI 12-JUL-2001
 LOCUS HOMO sapiens, RAB13, member RAS oncogene family, clone MGC:5074
 DEFINITION IMAGE:3451945, mRNA, complete cds.

ACCESSION BC000799
 VERSION BC000799.1 GI:12654002

KEYWORDS MGC;

SOURCE HOMO sapiens.

ORGANISM HOMO sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE 1 (bases 1 to 1211)

JOURNAL Strausberg, R.

Submitted (15-NOV-2000) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,

R. M.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 4 Row: 1 Column: 5

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 4506362.

Location/Qualifiers

1. 1211

/organism="Homo sapiens"

/db_xref="LocusID:5872"

/db_xref="taxon:9606"

/clones="MGC:5074 IMAGE:3451945"

/tissue_type="Placenta, choriocarcinoma"

/clone_lib="NH_MGC_10"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

95. .706

CDS

/codon_start=1
 /product="RAB13, member RAS oncogene family"
 /protein_id="AAH00799.1"
 /db_xref="GI:12654002"
 /translation="MAKAYDHLFKLLIGDSGVGKTCIIIRFAEDNNNTVISTIGID
 FKTRTVDEGKKIKLOVWDTPAGOEKFTITTVYRGAMGIIIVYDIDTDEKSFENIQNW
 MRSIKENASGVVERLLGNKCDMEARKVKQEKADKLAREHGIRFFETSAKSMNVDE
 AFSSLDRIILLKSGRRSGNPPSPSTDLTCDKKNNTKCSLG"
 BASE COUNT 392 a 256 c 335 g 228 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 8.38e-44 Length: 1211
 Score: 534.50 Matches: 102
 Percent Similarity: 66.38% Conservative: 54
 Best Local Similarity: 43.40% Mismatches: 55
 Query Match: 48.37% Indels: 25
 DB: 9 Gaps: 1

US-09-817-198B-2 (1-212) x BC000799 (1-1211)

QY 1 MetAlaLysGlnTyTrpValLeuPheArgLeuLeuIleGlyAspSerGlyValGly 20
 Db 95 ATGCCAAAGCCCTACGACCACTCTTCAAGTGTCTGCTGATCGGGGATCGGGGTGGC 154
 QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40
 Db 155 AAGACTTGTCTGATCATCTTGTTCAGAGGACAACTTCAACACACTTACATCTCCACC 214
 QY 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyLysValArgIleGln 60
 Db 215 ATCGGAATTGATTTCAAGATCCGCACTGTGATATAGGGGGAAGAGTCAAACTACA 274
 QY 61 IleTrpAspThrAlaGlyGlnGluArgTyTrpGlnThrIleThrLysGlnTyTrpArgArg 80
 Db 275 GTCTGGACACGGCTGCCAAGAGCGGTTCACAGACATAACTACTGCTACTACCGTGA 334
 QY 81 AlaGlnGlyIlePheLeuValTyTrpAspIleSerSerGluArgSerTyTrpGlnHisIleMet 100
 Db 335 GCCATGGGCATATTCTCTAGTATAGACATCCGATGAGAAATCTTTTCGAGAAATATTAG 394
 QY 101 LysTrpValSerAspValaspGluTyAlaProGluGlyValGlnLysIleLeuIleGly 120
 Db 395 AACTGGATGAAGACATCAAGGAGATGCCTCGCTGGGTGGAGCGCCCTCTCTCTGGGG 454
 QY 121 AsnLysAlaaspGluGluGlnLysArgGlnValGlyArgGlnGlnGlnGlnLeuAla 140
 Db 455 ACAAATGTGCATGGAGGCCAAGAGAGGTGCAGAGGAGCGGCGGATAGTGGCT 514
 QY 141 LysGluTyGlyMetAspPheTyTrpGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
 Db 515 CGAGAGCATGGAATCCGATTTTTCGAAACTAGTCTAAATCCAGTATGAATGTGGATGAG 574
 QY 161 SerPheThrArgLeuThrGluLeuValLeu----- 170
 Db 575 GCTTTTAGTTCCTGGCCGGGACATCTTCTCAAGTCAGGAGCGCGGATCAGGAAC 634
 QY 171 -----GlnAlaHisArgLysGlu 176
 Db 635 GGCACCAAGCCTCCAGTACTGACCTGAAACTTGTGCACAGAGAGAA-CACCAACAGTG 693
 QY 177 LeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGlu 196
 Db 694 CTCCTCGGTGAGGACCTTTCTTCTGCTCCCGCCCGGAGGTGAACCTGAGGAGAGAC 753
 QY 197 GluGlyLysProGluGlyProAlaAsnSerSerLysThrCysTrp 211
 Db 754 AACGGCAGAGGGAGTGCAGCGGGGAGAAATAGCAGAGGGGCTTGG 798
 RESULT 15
 HSRAB13
 LOCUS H. sapiens mRNA for rab 13.
 DEFINITION H. sapiens mRNA 1238 bp mRNA linear PRI 02-FEB-1994

[illegible]

Search completed: June 23, 2003, 11:52:37
Job time : 2491 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 23, 2003, 03:05:27 ; Search time 227 Seconds
(without alignments)
2103.188 Million cell updates/sec

Title: us-09-817-198b-2

Perfect score: 1105

Sequence: 1 MAKQDVLFRLLLIGSGVG.....LEEREGKPEGPANSSKTCWC 212

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-O=Cg2_1/USPTO.spool/US09817198/runat_18062003_144632_27901/app_query.fasta_1.391
-DB=N_Geneseq_101002 -QPMF=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09817198 -CGN_1_1_263 -runat_18062003_144632_27901 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_101002: *
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT: *
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	1105	100.0	2021	22	AAS27053	CDNA encoding nove
2	1092	98.8	1666	21	AAS75813	Human ORFX ORF1368
3	832	75.3	566	22	AAS27472	CDNA encoding nove
4	546.5	49.5	1540	23	ABL27707	Drosophila melanog
5	545	49.3	1986	22	ABA09160	Human rab8 homolog
6	540	48.9	2411	24	ABO54410	Human ovarian anti
7	534.5	48.4	1274	21	AAF18072	Lung cancer associ
8	530.5	48.0	716	21	AAA40104	Human Rab10 CDNA
9	530.5	48.0	861	21	AAA40108	Human Rab10 CDNA #
10	530.5	48.0	888	21	AAA40109	Canine Rab10 CDNA
11	530.5	48.0	956	21	AAA96887	Nucleotide sequenc
12	530.5	48.0	1537	21	AAF16196	Human prostate can
13	530.5	48.0	3533	22	AAH17889	Human CDNA sequenc
14	529.5	47.9	674	23	AAS71453	DNA encoding novel
15	528.5	47.8	1161	22	AHL75182	Nucleotide sequenc
16	528.5	47.8	2247	23	ABL29661	Drosophila melanog
17	528.5	47.8	2497	22	AAS60878	Human cancer agent
18	528.5	47.8	2497	22	AAS60884	Human cancer agent
19	528.5	47.8	2497	22	AAS60893	Human cancer agent
20	528.5	47.8	2497	22	AAS60895	Human cancer agent
21	528.5	47.8	2497	23	ABV25781	Human prostate exp
22	528.5	47.8	2497	23	ABV30037	Human prostate exp
23	528.5	47.8	3077	22	AAH13912	Human CDNA sequenc
24	519.5	47.0	866	22	AAH04301	Human CDNA clone (
25	519	47.0	1023	21	AAC51491	Arabidopsis thalia
26	519	47.0	1025	21	AAC34347	Arabidopsis thalia
27	513.5	46.5	1129	21	AAC44482	zea mays DNA fragm
28	510	46.2	911	21	AAC38429	Arabidopsis thalia
29	507	45.9	1101	21	AAC34080	Arabidopsis thalia
30	504	45.6	1203	21	AAC34385	zea mays DNA fragm
31	501	45.3	1705	21	AAC49208	Arabidopsis thalia
32	494	44.7	1153	21	AAC32781	Arabidopsis thalia
33	478.5	43.3	925	20	AAK27232	Arabidopsis thalia
34	478.5	43.3	939	22	AAK21661	Human Rab protein,
35	478.5	43.3	1898	22	AAS44924	Human breast and o
36	478.5	43.3	1944	22	AAI93456	CDNA encoding nove
37	477.5	43.2	730	21	AAC36696	Human polynucleoti
38	477.5	43.2	777	21	AAC42684	Nucleotide sequenc
39	477.5	43.2	881	24	ABN98761	Arabidopsis thalia
40	472.5	42.8	607	22	AAH87926	Arabidopsis thalia
41	472.5	42.8	1202	21	AAF14053	Peppermint plant o
42	468	42.4	959	21	AAC33987	Aspergillus oryzae
43	467	42.3	723	24	ABR83978	Arabidopsis thalia
44	467	42.3	2528	21	AAC60009	Human CDNA differe
45	466	42.2	932	21	AAC35200	Human secreted pro
						Arabidopsis thalia

ALIGNMENTS

RESULT 1

AAS27053

ID AAS27053 standard; CDNA; 2021 BP.

XX AAS27053;

XX 07-NOV-2001 (first entry)

XX cDNA encoding novel signal transduction pathway protein, Seq ID 88.

XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
XX antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
XX Immune system disorder; rheumatoid arthritis; inflammatory condition;
XX organ transplant rejection; infection; hepatitis C; blood disorder;
XX sickle cell anemia; hyperproliferative disorder; Gaucher's disease;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX chromosomal abnormality; Down syndrome; ischemia; renal disorder;
XX cardiovascular; respiratory; wound healing; endocrine; Addison's disease;

KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
KW acquired immune deficiency syndrome.

XX Homo sapiens.

OS WO200154733-A1.

PN PD

XX 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01312.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226686.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.

PR 08-SEP-2000; 2000US-0232081.

PR 12-SEP-2000; 2000US-0231968.

PR 14-SEP-2000; 2000US-0232397.

PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.

PR 14-SEP-2000; 2000US-0232400.

PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.

PR 14-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249266.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.

PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.

PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI: 2001-465460/50.

P-PSDB: AAU17555.

Novel polypeptides useful for diagnosing, treating, preventing and/or
prognosing disorders related to the proteins, including cancers, immune
disorders and neuronal disorders

Claim 1; SEQ ID No 507; 880pp; English.

The invention relates to novel isolated polypeptides (I), and
polynucleotides (II). (I), (II) and the antibody to (I) are useful for
diagnosing, preventing and treating diseases including immune system
disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
transplant rejections and graft versus host disease, infectious diseases
(e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
other blood-related disorders (sickle cell anaemia), myeloproliferative
disorders, primary haematopoietic disorders, hyperproliferative
disorders (e.g. Gaucher's disease and cancer), neurodegenerative
disorders (e.g. Alzheimer's disease, Parkinson's disease), renal
abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
disorders (e.g. glomerulonephritis), cardiovascular disorders
(e.g. arrhythmia), respiratory disorders, dermatological disorders, in
wound healing, epithelial cell proliferation, endocrine disorders (e.g.
Addison's disease), reproductive system disorders, gastrointestinal
disorder (inflammatory disorders), liver disorders (cirrhosis),
as stimulators of B-cell responsiveness to pathogens, activators of
T-cells, to induce higher affinity antibodies, and as a means to induce
tumour proliferation in pathologies e.g. acquired immune deficiency
syndrome (AIDS). AAS26976-AAS27850 represent novel signal transduction
pathway protein coding sequences and PCR primers of the invention.

Alignment Scores:

Pred. No.:	4.6e-87	Length:	566
Score:	832.00	Matches:	160
Percent Similarity:	97.56%	Conservative:	0
Best Local Similarity:	97.56%	Mismatches:	4
Query Match:	75.29%	Indels:	0
DB:	22	Gaps:	0

US-09-817-198B-2 (1-212) x AAS27472 (1-566)

Qy 1 MetalalysGlnTyrAspValLeuPheArgLeuLeuLeuGlyAspSerGlyValGly 20

Db 71 ATGCCGAAGCAGTACGATGCTGCTCGGCTGCTGTGATCGGGACTCCGGGTGGGC 130

XX	Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW	haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW	inhibin; chemotaxis; chemokines; thrombolysis; oncogenesis;
KW	proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW	myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW	chronic inflammatory condition; proliferative retinopathy;
KW	atherosclerosis; coronary heart disease; arterial ischaemia;
KW	bone disorder; osteoporosis; vascular growth disorder;
KW	tissue regeneration; wound healing; infection; immune disorder;
KW	cell culture; drug screening; gene therapy; antiinflammatory;
KW	antiarthritic; haemostatic; antiatherosclerotic;
KW	cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW	antifungal; vulvulnary; antiulcer; ss.

Homo sapiens.

WO200157188-A2.

09-AUG-2001.

05-FEB-2001; 2001WO-US03800.

03-FEB-2000; 2000US-0496914.

27-APR-2000; 2000US-0560875.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT;

WPI; 2001-457740/49.

P-PSDB; ABB11916.

Human proteins and DNA encoding sequences useful for preventing, treating or ameliorating a medical condition in a mammalian subject e.g. arthritis and cancer -

Claim 1; Page 805; 1963pp; English.

Sequences ABBI0981-ABBI2330 represent 1350 novel human polypeptides, and sequences ABA08225-ABA09574 represent nucleic acids encoding them. The invention also relates to vectors and recombinant host cells comprising a nucleotide of the invention, methods of producing the novel polypeptides, antibodies against the polypeptides, methods of detecting the nucleotides or polypeptides in a sample, and methods of identifying compounds which bind to polypeptides of the invention. Although novel, many of the polypeptides of the invention have homology to known proteins, thereby giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; haematopoiesis regulatory activity; tissue growth activity; immunomodulatory activity; activin- or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders. Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease or accidental damage. The polypeptides and nucleotides

198	Db	ACCATAGGAATAGACTTCCAGATCAAAACAGTAGTGAATTACAGGAAGAGATCAAGCTA	257
60	QY	GlnIleTtpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLyysGlnTyrTtrArg	79
258	Db	CAGATATGGGATACAGCAGGCCAGGCGATTCACACCATCACACCTCCTACTACAGA	317
80	QY	ArgAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIle	99
318	Db	GGCGCAATGGGTATCATCTGTTATATGACATCACCAAATGGTAAAGTTTGGAAACAATC	377
100	QY	MetLysTtpValSerAspValAspGluTyrAlaProGluGlyValGlnLyssIleLeuIle	119
378	Db	ACGCAATGGCTTGAAGAACATAGATGAGCATGCCAATGAGATGTGGAAAGAATGTACTA	437
120	QY	GlyAsnLysAlaAspGluGlnGlnLysArgGlnValGlyArgGlnGlnGlnGlnLeu	139
438	Db	GGAAACAAGTGTATATGGAGCACAAAGAGATTGTACCTTAACGGAAGAGAGACAGATT	497
140	QY	AlaLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLys	159
498	Db	GCAAGGGAGCATGGTATTAGTGTTTTTTGGAGCTAGTGCAAAACCAAAATATATAACATCGAA	557
160	QY	GluSerPheThrArgLeuThrGluLeuValLeuGln	171
558	Db	AAGGCGTTCCACAGTTAGCTGAGAGATATCCTTCGA	593

RESULTS

RESULT 9
AAA40108
ID AAA40108 standard; cDNA; 861 BP.

XX
AC
XX
DT

XX DE Human Rab10 cDNA #2. YY

AA Human; Rab10; Yap/Ral

XX	Homo sapiens.
OS	
XX	
PN	CN1249345-A.
XX	
XX	
PD	05-APR-2000.
XX	
XX	
PF	28-SEP-1998; 98CN-0121911.
XX	
XX	
PR	28-SEP-1998; 98CN-0121911.
XX	
XX	(UYFU-) UNIV FUDAN.
PA	
XX	
PI	Yu L, Tu Q, Gao J;
XX	
XX	
DR	WPI; 2000-400723/35.
XX	
PT	Preparation of human gene coding sequence, its encoded polypeptide -
XX	
XX	Disclosure; Fig 1; 23pp; Chinese.
PS	
XX	
CC	This invention describes a novel human Rab10 cDNA sequence. The protein
CC	coded by said sequence belongs to Yap/Rab family and is the homolog of
CC	mouse Rab10. The present invention also relates to the polypeptide coded
CC	by said nucleotide sequence and the application and preparing process
CC	of said polynucleotide and said polypeptide. This sequence encodes the
CC	human Rab10 protein described in the method of the invention.

DB:	21	Gaps:	1
 US-09-817-198B-2 (1-212) x AAAA40108 (1-861)			
QY	1	MetAlaLysGln--TyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyVal	19
Dd	95	ATGGCGAAAGACGCTAGCAGCTGCTTTCAAGTGCTGTGATCGGGGATTCCGGAGTG	154
QY	20	GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisIleSer	39
Dd	155	GGGAAGACCTGGTCCTTTTTTCGTTTTTCGGATGATGCCCTTCAACTACTACCTTTATTTC	214
QY	40	ThrIleGlyValAspPheLysMetLysThrIleGluValAspGlyLysValArgIle	59
Dd	215	ACCATTAGGAATAAGATTCCAAGATCAAACAGCTTCAATTACAAGAGAAGACATCAAGCTA	274
QY	60	GlnIleTrpAspThrAlaGlyGlnGluArgTyrrGlnThrIleThrLysGlnTyrrArg	79
Dd	275	CAGATTGGGATACAGCAGGCAGGCGGATTTCCACCATCACACCTCCTCTACTACAGA	334
QY	80	ArgAlaGlnGlnIlePheLeuValTyrrAspIleSerSerGluArgSerrTyrrGlnHisIle	99
Dd	335	GGCGCAATGGGTATCATCTAGTAGTATGACATCACCAATGGTAAGATTTTGAACAATC	394
QY	100	MetLysTrpValSerAspValAspGluTyrrAlaProGluGlyValGlnLysIleLeuIle	119
Dd	395	AGCAAAATGGCTTAGAAACATAGATGAGCATGCCAATGAAGATGTGGAAGAATGTTACTA	454
QY	120	GlyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeu	139
Dd	455	GGAAACAAGTGTGATATGGCGACAAAAGAGTCTTCTCTAAAGGAAAAAGGAGAACAGAT	514
QY	140	AlaLysGluTyrrGlyMetAspPheTyrrGluThrSerAlaCysThrAsnLeuAsnIleLys	159
Dd	515	GCAAGGGAGCATGGTATTAGTGTTTTTTGAGACTACTGCAAAAGCAAAATATAAACATCGAA	574
QY	160	GluSerPheThrArgLeuThrGluLeuValLeuGln	171
Dd	575	AAGGGGTTCCTCAGCTTAGCTGAGATATCCTTCGA	610

Preparation of human gene coding sequence, its encoded polypeptide -
 XX
 XX
 PS Disclosure; Fig 1; 23pp; Chinese.
 XX
 XX
 CC This invention describes a novel human Rab10 cDNA sequence. The prot

CC coded by said sequence belongs to Yap/Rab family and is the homolog of
CC mouse Rab10. The present invention also relates to the polypeptide coded
CC by said nucleotide sequence and the application and preparing process
CC of said polynucleotide and said polypeptide. This sequence encodes a
CC canine Rab10 protein described in the method of the invention.

Sequence 888 BP; 261 A; 202 C; 202 G; 223 T; 0 other;
SQ

Alignment Scores:

Pred. No.:	7.38e-52	Length:	888
Score:	530.50	Matches:	96
Percent Similarity:	82.56%	Conservative:	46
Best Local Similarity:	55.81%	Mismatches:	29
Query Match:	48.01%	Indels:	1
DB:	21	Gaps:	1

US-09-817-198B-2 (1-212) x AAA40109 (1-888)

QY	1	MetAlaLysGln---TyrAspValLeuPheArgLeuLeuLeuGlyAspSerGlyVal	19
Db	133	ATGGCGAAGAACACGTAGACGCTGTTTTCAAGCTGCTCTGATCGAGAGATCGGGAGTA	192
QY	20	GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSer	39
Db	193	GGGAAGACCTGGCTCTCTTTTCGTTTTTCGGATGATGCCTTCATACCACTTTATTATTC	252
QY	40	ThrIleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIle	59
Db	253	ACCATAGGAATAGATTTTAAGATCAAAACAGCTTGAATTACAAGGAAGAAGATCAAGCTA	312
QY	60	GlnIleTrpAspThrAlaGlyGlnClnuArgThrGlnThrIleThrLysGlnTyrTyrArg	79
Db	313	CAGATATGGGATACAGCAGGCGAGGCGATTTCCACCATCACAACTCCTACTACAGA	372
QY	80	ArgAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIle	99
Db	373	GGAGCAATGGGTATCATGCTAGTATATGACATACCACTGGTAAAAAGCTTTTGAAAAACATC	432
QY	100	MetLysTrpValSerAspValAspClnuTyrAlaProGluGlyValGlnLysIleLeuIle	119
Db	433	AGCAATATGGCTTAGAAACATAGATGAGATGCCAATGAAGATGGGAAGAAGTTGTTACTA	492
QY	120	GlyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLysLeu	139
Db	493	GGAAACAAATGTGATATGGACGATAAAAGAGTTGTCTACCTAAAGGAAAAAGGAGACAGATT	552
QY	140	AlaLysGluTyrGlyMetAspPheTyrClnuThrSerAlaCysThrAsnLeuAsnIleLys	159
Db	553	GCAGGGAGCATGGTATATAGATTTTTCGAGACTAGTGCAAAAGTAAATATAAACAATCGAA	612
QY	160	GluSerPheThrArgLeuThrGluLeuValLeuGln	171
Db	613	AAGGGTTTCCTCACATATTAGCTGAAGATATCCTCGA	648

RESULT 11

AAA96887
ID AAA96887 standard; DNA: 956 BP.

AA
AC
AAA96887;

XX
DT 19-FEB-2001 (first entry)

XX Nucleotide sequence of human RAB10.
DE

XX
KW
RAB protein; GTPase; GTP binding; gene therapy; cancer; ss.

XX
OS Homo sapiens.

XX	Key	Location/Qualifiers
FH	CDS	127..729
FT		/+tag= a
FT		/product= "FAB10"
FT		

Db
487 GGAAACAAGTGTGATATGGACGACAAAAGAGTTGTACCTAAAGGAAAAGGAGAACAGATT 546

WO200058464-A2.

XX
XX

PD 05-OCT-2000.

XX

PF 13-MAR-2000; 2000WO-US0

XX
PR 25-MAR-1999; 99US-0126083.
XX
PA (AXYS-) AXYS PHARM INC.

XX
XX
XX

PI Allen M, Abel K, McIntosh B, Vega R, Rutter M, Buckler AJ;
 DR WPI; 2000-647233/62.
 DR P-PSDB; AAB19165.
 XX
 XX
 PT Novel isolated nucleic acid encoding a mammalian RAB protein useful for
 PT identifying homologous or related genes, in producing composition that
 PT modulates expression or function of RAB for cancer therapy -

PS Claim 4; Page 38-39; 58pp; English.

The present sequence encodes a mammalian RAB protein. RAB proteins constitute the largest family of small GTPases, with over 40 identified isoforms. RAB proteins contain four highly conserved peptide sequences involved in GTP binding and hydrolysis. Compositions comprising RAB nucleic acid are useful for identifying homologous or related genes, in producing compositions that modulate the expression or function of RAB, for gene therapy, mapping functional regions of the protein and in studying associated physiological pathways. In addition, modulation of the gene activity *in vivo* is used for prophylactic and therapeutic purposes, such as treatment of cancer, and identification of cell type based on expression. The DNA may also be used to identify expression of the gene in a biological specimen.

Sequence 956 BP; 294 A: 205 C: 213 G: 244 T: 0 other:
AA
SQ

Alignment Scores:

Pred. No.:	8.16e-52	Length:	956
Score:	530.50	Matches:	96
Percent Similarity:	82.56%	Conservative:	46
Best Local Similarity:	55.81%	Mismatches:	29
Query Match:	48.01%	Indels:	1
DB:	21	Gaps:	1

US-09-817-198B-2 (1-212) X AAA96887 (1-956)

Oy 1 MetalaLysGln---TyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerClyVal 19
 |||||
Dd 127 ATGC CGAAGAAGCTAC GACCCTTTTC AAGTGCTCTGATCGGGGATTCCGGAGTG 186

Qy	20	Gly	Thr	Cys	Leu	Cys	Arg	Phe	Thr	Asn	Glu	Phe	His	Ser	Ser	His	Leu	Ser	39
Db	187	GGA	AGAC	TGC	GCT	TTT	TGG	TAT	TTC	TTT	TGG	GAT	GC	CTT	CA	ACT	ACT	TAT	TTC

Oy 40 ThrileGlyValaspPhelysmethylstThrileGlualaspGlylelYsvAlargile 59
 | ||| | |||| | | | | | | | | | | | | | |
Dh 247 ACCTATAGCAATACACTTTTAAAGATCAAACAACCTTCATAAACAACAGCATCACCAGA 205

Qy 60 GlnIleTspAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArg 79

Qy 80 ArgAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIle 99

Oy	100	MethylstypvalSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIle	119
----	-----	---------------------------------------------------------------	-----

DB	QY	120	421
DD	QY	120	421

QY 140 AlaLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLys 159
 DB 547 GCAGGGAGCAGTGGTATTAGGTTTTCAGACTAGTGCAGAAAGCAATATATAACATCGAA 606
 QY 160 GluSerPheThrArgLeuThrGluLeuValLeuGln 171
 DB 607 AAGGGTTCCTCACGTTAGCTGAGATATCTTCGA 642

RESULT 12

ID AAF16196 standard; cDNA; 1537 BP.
 AC AAF16196;
 XX
 XX
 DT 13-MAR-2001 (first entry)

Human prostate cancer antigen nucleotide sequence SEQ ID NO:631.

DE Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
 KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
 KW vulnery; gastrointestinal; nephrotropic; antiinfective; gynaecological;
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
 KW wound; infectious disease; ss.

OS Homo sapiens.

PN WO200055174-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05988.

PR 12-MAR-1999; 98US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

PI Rosen CA, Ruben SM;

XX WPI; 2000-587513/55.

DR P-PSDB; AAB56993.

PT Prostate cancer associated gene sequences, referred to as prostate
 PT cancer antigens, useful for treatment, prevention, and diagnosis of
 PT disorders such as prostate cancer -

PS Claim 1; Page 1073; 2338pp; English.

CC AAF15566 to AAF16505 encode the human prostate cancer associated
 CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
 CC The prostate cancer antigens can have neuroprotective, cytostatic,
 CC cardioactive, immunomodulatory, muscular, vulnery, gastrointestinal,
 CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
 CC and can be used in gene therapy. The prostate cancer antigen
 CC polynucleotides may be used for detection of prostate cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The prostate cancer antigens may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
 CC AAB57303 represent sequences used in the exemplification of the present
 CC invention.

SQ Sequence 1537 BP; 405 A; 378 C; 384 G; 369 T; 1 other;

Alignment Scores:

Pred. No.: 1537
 Score: 1.56e-51 Length: 1537
 530.50 Matches: 96
 Percent Similarity: 82.56% Conservative: 46
 Best Local Similarity: 55.81% Mismatches: 29
 Query Match: 48.01% Indels: 1

DB: 21 Gaps: 1
 US-09-817-198B-2 (1-212) x AAF16196 (1-1537)
 QY 1 MetalAlaLysGln---TyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyVal 19
 DB 533 ATGGCGAAGAGACGCTACGACCTGCTTTTCAGCTGCTCCTGATCGGGATTCGGAGTG 592
 QY 20 GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSer 39
 DB 593 GGGAAAGACCTGCTCTTTTCGTTTTCGGATGATGCTTCAATACCTATACCTTTATTC 652
 QY 40 ThrIleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIle 59
 DB 653 ACCATAGGAATAGACTTCAAGATCAAAACAGTTGAATTACAGAGAAAGATCAAGCTA 712
 QY 60 GlnIleTyrAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrArg 79
 DB 713 CAGATATGGGATACAGCAGGCCAGGCGGATTTACACCATCACAACCTCTACTACAGA 772
 QY 80 ArgAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIle 99
 DB 773 GCGCAATGGGTATCATCTAGTATATGACATCACCAATGGTAAAGTTTGAACATC 832
 QY 100 MetLysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIle 119
 DB 833 AGCAAAATGGCTTAGAAACATAGATGAGCATGCCAATGAAGATGTGCAAGAATGTACTA 892
 QY 120 GlyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLys 139
 DB 893 GGAACAAGTGTGATATGGACGACAAAGAGTTGTACCTAAGGAAAAGGAGACAGATT 952
 QY 140 AlaLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLys 159
 DB 953 GCAAGGGAGCATGGTATTAGGTTTTCAGACTAGTGCAGAAAGCAATATAACATCGAA 1012
 QY 160 GluSerPheThrArgLeuThrGluLeuValLeuGln 171
 DB 1013 AAGCGTTCTCCTACGTTAGCTGAGATATCTTCGA 1048
 RESULT 13
 AAF17889
 ID AAF17889 standard; cDNA; 3533 BP.
 XX
 AC AAF17889;
 XX
 XX 26-JUN-2001 (first entry)
 XX Human cDNA sequence SEQ ID NO:17618.
 DE
 DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 KW Homo sapiens.
 OS
 XX EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 XX 28-JUL-2000; 2000EP-0116126.
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX

PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX
PS
PS
XX
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 3533 BP; 967 A; 716 C; 730 G; 1120 T; 0 other;

Alignment Scores:
Pred. No.: 4, 85e-51 Length: 3533
Score: 530.50 Matches: 96
Percent Similarity: 82.56% Conservative: 46
Best Local Similarity: 55.81% Mismatches: 29
Query Match: 48.01% Indels: 1
DB: 22 Gaps: 1

US-09-817-198B-2 (1-212) x AAH17889 (1-3533)

QY 1 MetAlaLysGln---TyrAspValLeuPheArgLeuLeuLeuLeuLeuLeuVal 19
DB 500 ATGGCGAAGAGACGACGACCTCTTCAAGCTGCTCTGATGGGATTCGGAGTG 559
QY 20 GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisIleSer 39
DB 560 GGAAGACCTGCTCTCTTTTCGTTTCGGATGATCCCTCACTACTATTATTC 619
QY 40 ThrIleGlyValAspPheLysMetLysThrIleGluValAspGlyLysValArgIle 59
DB 620 ACCATAGGAATAGACTCAAGATCAAAACAGATTGAATTACAGGAAGAGATCAAGCTA 679
QY 60 GlnIleTrpAspPheAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrArg 79
DB 680 CAGATATGGATACAGACGACGACGACGATTCACACCATCAACCTCTACTACAGA 739
QY 80 ArgAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIle 99
DB 740 GCGCGCAATGGTATCATCTAGTATGATGATGATGATGATGATGATGATGATGATGAT 799
QY 100 MetLysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeu 119
DB 800 AGCAAAATGGCTTGAAGAACATAGATGACATGCCAATGAAGATGTGGAAGAATGTACTA 859
QY 120 GlyAsnLysAlaAspGluGlnGluLysArgGlnValGlyArgGluGlnGlnGlnLeu 139
DB 860 GGAACAAGTGTGATATGGACGACAAAGAGTTGTATCCTTAAGGAAGAGAGACAGATT 919
QY 140 AlaLysGlnTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLys 159
DB 920 GCAAGGAGCATGTTATGTTTGTGACTAGTGTGCAAAACAAATATATAACATCGAA 979

QY 160 GluSerPheThrArgLeuThrGluLeuValLeuGln 171
DB 980 AAGCGCTTCTCACGTAGCTAGTGAAGATACCTTCGA 1015

RESULT 14

AA571453
ID AAS71453 standard; cDNA; 674 BP.

XX AAS71453;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #7257.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG07266.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX

PS Claim 1; SEQ ID No 7257; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 674 BP; 218 A; 136 C; 173 G; 147 T; 0 other;

Alignment Scores:

Pred. No.: 6, 62e-52 Length: 674
Score: 529.50 Matches: 96
Percent Similarity: 82.56% Conservative: 46
Best Local Similarity: 55.81% Mismatches: 29
Query Match: 47.92% Indels: 1

DB: 23 Gaps: 1

US-09-817-198B-2 (1-212) x AAS71453 (1-674)

QY 1 MetAlaLysGln---TyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyVal 19
Db | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
75 ATGCCGAAGAAGACGTACCACCTGGTTTCCTCAAGCTGCCTCCTCATCGGGATCCGGAGTG 134
QY 20 GlyLysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisLeuSer 39
Db | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
135 GGGAAGACTGGCTCTTTTCGTTTCGGATGATGCTTCATACACCTATTATTTCC 194
QY 40 ThrIleGlyValAspPheLysMetLysThrIleGluValAspGlyIleValArgile 59
Db | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
195 ACCATAGGAATAGACTTCAAGATCAAACAGTGAATTAACAAGGAAGAACAGCAAGCTA 254
QY 60 GluIleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTrg 79
Db | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
255 CAGATATGGATGACAGCAGCGGAGGAGGATTTTCACACCATCACACCTCTACTACAGA 314
QY 80 ArgAlaGlnGlyIlePheLeuValTyraSpilleSerSerGluArgSerTyrGlnHisle 99
Db | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
315 GGGCAATGGATCATCATCTAGTATATGACATCACCATCATGTTAAAGTTTGAACACATC 374
QY 100 MetLysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeu 119
Db | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
375 AGCAATGGCTTTATAACATAGATGAGCATGCAATGAAGATGTGGAAGAATGTACTA 434
QY 120 GlyAsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlyGlnLeu 139
Db | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
435 GGAACACATGTGATGATGACGACACAAAGAGTTGTACCTAAAGGAAGGAGAACAGATT 494
QY 140 AlaLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLys 159
Db | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
495 GCAAGGAGCATGGTATTAGTTTITTTGAGACTACTGCAAAAGCAAAATATAACATCGAA 554
QY 160 GluSerPheThrArgLeuThrGluLeuValLeuGln 171
Db | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
555 AAGCGCTTCTCAGCTAGCTGAAGATATCCTCTCGA 590

RESULT 15

AAH75182 standard; DNA; 1161 BP.

XX AAH75182;
AC | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
DT | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
XX 13-NOV-2001 (first entry)
DE Nucleotide sequence of human 27423 G-protein.
KW Human; G-protein; 32705; 23224; 27423; 32700; 32712; lung disorder;
KW congenital anomaly; pulmonary congestion; oedema; haemorrhage;
KW adult respiratory distress syndrome; Goodpasture's syndrome;
KW chronic obstructive pulmonary disease; asthma; pulmonary hypertension;
KW liver disorder; hepatic injury; jaundice; cholestasis; viral hepatitis;
KW cirrhosis; Wilson's disease; autoimmune hepatitis; hepatic failure;
KW brain disorder; hypoxia; cerebral ischemia; intracranial haemorrhage;
KW acute meningitis; Parkinson's disease; Alzheimer's disease; glioma;
KW chronic bacterial meningoencephalitis; multiple sclerosis;
KW amyotrophic lateral sclerosis; stroke; Huntington's disease; ss.
XX Homo sapiens.
OS
XX
XX Location/Qualifiers
FH 18..641
FT CDS /tag= a
FT /product= "G-protein"
XX
XX WO200164887-A2.
PN
XX
XX
PD
XX
XX 07-SEP-2001.
PF 27-FEB-2001; 2001WO-US06292.

29-FEB-2000; 2000US-0185606.
(MILL-) MILLENNIUM PHARM INC.
Meyers RA;
WPI; 2001-550182/61.
P-PSDB; AAG67154.
Novel human small G-protein polypeptides and polynucleotides for treating lung disorders, liver disorders and brain disorders -
Claim 2; Fig 15; 151pp; English.
The present sequence encodes a human G-protein. The specification describes 32705, 23224, 27423, 32700 or 32712 small G-proteins. The G-protein polypeptides and polynucleotides are useful as a target for diagnosis and treatment of G-protein mediated or related disorders, and for identifying agonists and antagonists for diagnosis and treatment. They are useful for treating disorders of lung (e.g. congenital anomalies, pulmonary congestion, oedema, adult respiratory distress syndrome, haemorrhage, chronic obstructive pulmonary disease, asthma, Goodpasture's syndrome and pulmonary hypertension), liver (e.g. hepatic injury, jaundice, cholestasis, viral hepatitis, cirrhosis, Wilson's disease, autoimmune hepatitis and hepatic failure), and brain (e.g. hypoxia, cerebral ischemia, intracranial haemorrhage, acute meningitis, Parkinson's disease, Alzheimer's disease, gliomas, chronic bacterial meningoencephalitis, multiple sclerosis, amyotrophic lateral sclerosis, stroke and Huntington's disease).
SQ Sequence 1161 BP; 368 A; 246 C; 243 G; 304 T; 0 other;

Alignment Scores:
Pred. No.: 1,81e-51 Length: 1161
Score: 528.50 Matches: 95
Percent Similarity: 80.68% Conservative: 47
Best Local Similarity: 53.98% Mismatches: 33
Query Match: 47.83% Indels: 1
DB: 22

US-09-817-198B-2 (1-212) x AAH75182 (1-1161)

QY 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValGly 20
Db | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
18 ATGCCGAAGACGTACGATTATCTCTCAAGCTCCTGCTGATCGGACATCGGGGTAGGC 77
QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisLeuSerThr 40
Db | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
78 AAGACCTGCCTCCTGTTCCGCTTCTCAGAGGACGCCCTTCAACACACCTTCATCTCCACC 137
QY 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgileGln 60
Db | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
138 ATCGGAATGATTTTAAAATTTAGAACGATAGACTAGATGGAAGAAATTAAGCTTCAG 197
QY 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTrgArg 80
Db | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
198 ATATGGGACACACGGGCTCAGGAAGATTCGGAACATCCGACACGCTACTACAGAGA 257
QY 81 AlaGlnGlyIlePheLeuValTyraSpilleSerSerGluArgSerTyrGlnHisleMet 100
Db | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
258 GCCATGGCATATTATGCTGCTATGACATCAAAATGAAAAATCTTTTGACAATATATAA 317
QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuGly 120
Db | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
318 AATTGGATCAGAAACAACTTGAAGAGCATGCTCTCCGATGTCCGAAGAAATGATCTGGGT 377
QY 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlyGlnLeuAla 140
Db | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
378 AACAAATGTGATATGATGAATGACAAAGACAAAGTGTCAAAGAAAGAGGAGAGTAGCA 437
QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
Db | | | | | : : : : : | | | | | : : : : : | | | | | : : : : : | | | | | : : : : :

Db 438 ATTGACTATGGATTAAATCTTGGAGACAAGCGCAAAATCCAGTCCAATGTAGAGAG 497

Qy 161 Ser---PheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLys 175

Db 498 GCATTTTTCACACTTGCACGAGATATATGACAAAACTCAACAGAAAA 545

Search completed: June 23, 2003, 11:10:07
Job time : 232 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 23, 2003, 10:52:29 ; Search time 1383 Seconds
(without alignments)
2482.605 Million cell updates/sec

Title: US-09-817-198b-2

Perfect score: 1105

Sequence: 1 MAKQDVLFRLLLIGDSGVG.....LEEEKGPEGPANSSKTWC 212

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1615406 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q/cgn2_1/USPTO.spool/US09817198/runat_18062003_144633_27919/app_query.fasta_1.391
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=0.1 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09817198.ecgn_1_1906_@runat_18062003_144633_27919 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: em_estba:
2: em_esthum:
3: em_estin:
4: em_estmu:
5: em_estov:
6: em_estpl:
7: em_estro:
8: em_hic:
9: gb_est1:
10: gb_est2:
11: gb_hic:
12: gb_est3:
13: gb_est4:
14: gb_est5:
15: em_estfun:
16: em_estom:
17: gb_gss:
18: em_gss_hum:
19: em_gss_inv:
20: em_gss_pln:
21: em_gss_vrt:
22: em_gss_fun:
23: em_gss_mam:
24: em_gss_mus:
25: em_gss_other:
26: em_gss_pro:
27: em_gss_rod:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	996	90.1	1091	12	BF160330	BF160330 601768601
2	953.5	86.3	932	12	BF535642	BF535642 602054039
3	922	83.4	616	13	BI648588	BI648588 603277781
4	914	82.7	992	12	BF101730	BF101730 601753464
5	910	82.4	902	12	BF178163	BF178163 601809083
6	885	80.1	699	13	BI557933	BI557933 603237549
7	858.5	77.7	796	12	BF966292	BF966292 602286692
8	838	75.8	1100	12	BF181167	BF181167 601805830
9	829	75.0	616	12	BG247902	BG247902 602359819
10	816	73.8	1121	12	BF182001	BF182001 601805936
11	799	72.3	814	12	BG762967	BG762967 602735022
12	792	71.7	660	13	BG974411	BG974411 602844184
13	777	70.3	911	13	BI649317	BI649317 603278051
14	756	68.4	618	12	BF150465	BF150465 uy86905.y
15	733	66.3	1767	13	BM462147	BM462147 AGENCOURT
16	727	65.8	485	12	BE913096	BE913096 601668580
17	706	63.9	667	13	BI685808	BI685808 603309467
18	704	63.7	843	12	BG865609	BG865609 602783763
19	693.5	62.8	577	12	BG703353	BG703353 602685115
20	686	62.1	465	12	BF565114	BF565114 UT-R-B01-
21	670	60.6	479	12	BF320679	BF320679 uz55d10.y
22	666	60.3	543	12	RG019552	RG019552 daa21e04..
23	663	60.0	539	12	BF022618	BF022618 uy51f07.y
24	658.5	59.6	693	10	BE287683	BE287683 601093322
25	651	58.9	823	13	BM017178	BM017178 603643780
26	649	58.7	385	14	BM936454	BM936454 UT-M-BH2.
27	589	53.3	956	14	BQ944882	BQ944882 AGENCOURT
28	568.5	51.4	704	13	BJ057892	BJ057892 B057692
29	554.5	50.2	714	13	BI259835	BI259835 602971485
30	551.5	49.9	644	14	BQ387194	BQ387194 NISC.Jm22
31	547	49.5	611	13	BJ035681	BJ035681 BJO35681
32	547	49.5	678	13	BI395002	BI395002 pcp1p.pk0
33	547	49.5	788	9	AU125167	AU125167 AU125167
34	545	49.3	613	14	BM791355	BM791355 K-EST0071
35	545	49.3	635	9	AU134059	AU134059 AU134059
36	545	49.3	676	9	AU133710	AU133710 AU133710
37	545	49.3	686	12	BG709122	BG709122 602675387
38	545	49.3	816	9	AU119700	AU119700 AU119700
39	545	49.3	835	12	BG425680	BG425680 602452987
40	545	49.3	860	9	AU122740	AU122740 AU122740
41	545	49.3	877	13	BI763900	BI763900 603049666
42	545	49.3	897	14	BQ671010	BQ671010 AGENCOURT
43	545	49.3	986	13	BM475359	BM475359 AGENCOURT
44	545	49.3	1032	13	BM459427	BM459427 AGENCOURT
45	545	49.3	1059	14	BM923211	BM923211 AGENCOURT

ALIGNMENTS

RESULT 1

BF160330

LOCUS

DEFINITION

BF160330

ACCESION

BF160330

VERSION

BF160330.1

KEYWORDS

EST.

SOURCE

ORGANISM

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 1091)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

BF160330 1091 bp mRNA linear EST 30-OCT-2000
601768601F1 NCI_CGAP_Lu29 Mus musculus cdna clone IMAGE:3987745 5',
mRNA sequence.

BF160330

BF160330.1

GI:11040541

EST.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 1091)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Qy 61 IletRspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
 Db 325 ATTGGGACACAGCAGGCGAGGAGGTACCAAGCAGTACTATCGCGCA 384
 Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
 Db 385 GCCCGGGAATATTTTGTCTAGCAGATTTAGCAGTGGCGCTCTATCATCATATCATG 444
 Qy 101 LysTrpValSerAspValaspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
 Db 445 AAGTGGGTCACTGACGTGGATGAGTACGCTCCAGAGGAGTCCAGAAAGATCCTAATGGG 504
 Qy 121 AsnLysAlaaspGluGlnGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
 Db 505 AATAAGGCTGATGAAGACAGAAACGGCAGTGGGAGAGAGCAGGCGGAGCTGGCT 564
 Qy 141 LysGlnTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
 Db 565 AAGGATGACGCTGAGTGGTCTTACGAAACAGTGGCTGGCAGACACCTTATATTCAGAG 624
 Qy 161 SerPheThrArgLeuThrGluLeuValLeuAlaHisArgLysGluLeuGluGlyLeu 180
 Db 625 TCCTTCACTGCTGACGGAGCTGGTGTCTGAGGGCCACAGGAAAGAGCTGGATGGGCTC 684
 Qy 181 ArgMetArgAla-SerAsnGlnLeuAlaLeuAla---GluLeuGluGluGluGlyLy 199
 Db 685 CGAACAGTCCCGCGCAACAGCTCGCACATGGCGGAGACTGTGAGACGCGCAAGCGA 744
 Qy 199 sProGluGlyPro 203
 Db 745 ACCGTATGGGCCA 757

RESULT 3
 BI648588
 LOCUS 603277781F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5317917 5',
 DEFINITION mRNA sequence.
 ACCSSION BI648588
 VERSION BI648588.1 GI:15562824
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLNL1806 row: a column: 22
 High quality sequence start: 2
 High quality sequence stop: 614.
 Location/Qualifiers
 1. 616
 /organism="Mus musculus"
 /strain="129,C57BL/6J,FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5317917"
 /clone_lib="NCI_CGAP_Mam3"
 /tissue_type="tumor, gross tissue"
 /dev_stage="10 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
 Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
 Reference for transgenic model: Xu et al., Nature Genetics
 22, 37-43 (1999)."

BASE COUNT 167 a 154 c 179 g 115 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 8.89e-108 Length: 616
 Score: 922.00 Matches: 180
 Percent Similarity: 98.92% Conservative: 3
 Best Local Similarity: 97.30% Mismatches: 2
 Query Match: 83.44% Indels: 1
 DB: 13 Gaps: 0

US-09-817-198b-2 (1-212) x BI648588 (1-616)

Qy 28 PheThrAspAsnGlnPheHisSerSerHisIleSerThrIleGlyValaspPheLysMet 47
 Db 25 TTCACCGACACAGAGTTCCTCTCGCATATCTCCACCATCGGTGTGATTTAAGATG 84
 Qy 48 LysThrIleGluValaspGlyIleLysValArgIleGlnIleTrpAspThrAlaGlyCln 67
 Db 85 AAGACTATCGATGTAGACGGCATCAAGTGAGATACAGATTGGGACACAGAGGCGAG 144
 Qy 68 GluArgTyrGlnThrIleThrLysGlnTyrTyrArgArgAlaGlnGlyIlePheLeuVal 87
 Db 145 GAGAGGTACCAAGACTATCACAAAGCAGTACTATCGGCGAGCCCA-GGAATATTTTATGTC 203
 Qy 88 TyrAspIleSerSerGluArgSerTyrGlnHisIleMetLysTrpValSerAspValasp 107
 Db 204 TACGACATTTAGCAGTGGCTCTCTATCAGCATATCATGAAGTGGGTGAGTGGAT 263
 Qy 108 GluTyrAlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAlaaspGluGln 127
 Db 264 GAGTACGCTCCAGAAAGAGTCCAGAGATCTCTAATTGGATTAAGCTGATGAAGAGCAG 323
 Qy 128 LysArgGlnValGlyArgGluGlnGlyClnGlnLeuAlaLysGluTyrGlyMetaspPhe 147
 Db 324 AAACGGCAGTGGGGAGAGAGCAGGCGCAGCTGGCTTAAGAGTAGTACGGCATGGACATC 383
 Qy 148 TyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThrArgLeuThrGlu 167
 Db 384 TACGAAACAGTGGCTCGACCACTTATATTAAGAGTCTCTACCTGCTGACGGAG 443
 Qy 168 LeuValLeuGlnAlaHisArgLysGluLeuGluGlyLysProGluGlyProAlaAsnSerSer 187
 Db 444 CTGGTGTGCGAGGCCACAGGAAAGAGCTGGTCTCCGACACAGCTGCCCAACAGCAG 503
 Qy 188 LeuAlaLeuAlaGluLeuGluGluGluGlyLysProGluGlyProAlaAsnSerSer 207
 Db 504 CTCGCACTGGCGGAGCTGGAGGAGGAGGAGGCAAACTGAGGGCCCACTCTTCA 563
 Qy 208 LysThrCysTrpCys 212
 Db 564 AAGACCTGCTGGTGC 578

RESULT 4

BI6101730 992 bp mRNA linear EST 19-OCT-2000
 LOCUS 601753464F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3981183 5',
 DEFINITION mRNA sequence.

ACCESSION BI6101730

VERSION BI6101730.1 GI:10884256

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NIH-MGC http://mgi.nci.nih.gov/.

1 (bases 1 to 992)

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Qy 70 TyrGlnThrIleThrLysGlnTyrTyrArgAlaGlnGlyIlePheLeuValTyrAsp 89
 Db 134 TACAGACTATCAAAAGCAGTACTATCGCGAGCCAGGGAATATTTTAGTCTACGAC 193
 Qy 90 IleSerSerGluArgSerTyrGlnHisIleMetLysTyrValSerAspValAspGluTyr 109
 Db 194 ATTAGCAGTACGCGCTCTATCAGCATATCATGAAGTGGTCACTGAGTGATGATAC 253
 Qy 110 AlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAlaAspGluGlnLysArg 129
 Db 254 GCTCCAGAGAGTCCAGAGATCCTAATTGGGAATTAAGCTGATGAAGAGCAAAACGG 313
 Qy 130 GlnValGlyArgGluGlnGlnGlnLeuAlaLys-GluTyrGlyMetAspPheTyrGI 149
 Db 314 CAGGTGGGAGAGAGCAGGAGGAGCAGCTGGCTAACGGATACGGATGAGCTTCTACGA 373
 Qy 149 uThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThrArgLeuThrGluLeuVa 169
 Db 374 AACAAAGTGCCTGCACCAACCTTAATATTAAGAGTCTTTCACCTGCTGACGGAGCTGGT 433
 Qy 169 IleGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAl 189
 Db 434 GCTCAGGCCACAGGACAGAGTGGATGGTCTCCGAACACAGTGCACCAAGAGCTCGC 493
 Qy 189 aLeuAlaGluLeuGluGluGluGlyLysProGluGlyProAlaAsnSerSerLysTh 209
 Db 494 ACTGGCCGAGCTGGAGGAGGAGCAAGCAACCTTGAGGGCCACGAAACTCTTCAAAAGC 553
 Qy 209 rCysTyrCys 212
 Db 554 CTGCTGGTGC 563

RESULT 6
 BI557933
 LOCUS 603237549f1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5290284 5',
 DEFINITION 699 bp mRNA linear EST 05-SEP-2001
 mRNA sequence.
 ACCSSION BI557933
 VERSION BI557933.1 GI:15445247
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 699)
 TITLE NIH-MGC http://mgi.nci.nih.gov/.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
 cDNA Library Prepared by: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM11734 row: b column: 13
 High quality sequence stop: 698.
 Location/Qualifiers
 1..699
 /organism="Mus musculus"
 /strain="129,C57BL/6J,FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5290284"
 /clone_lib="NCI_CGAP_Mam3"
 /tissue_type="tumor, gross tissue"
 /dev_stage="10 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Sali;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH

Reference for transgenic model: Xu et al., Nature Genetics
 22, 37-43 (1999)."

BASE COUNT 164 a 188 c 128 t
 ORIGIN

Alignment Scores:
 Pred. No.: 6.le-103 Length: 699
 Score: 885.00 Matches: 172
 Percent Similarity: 98.87% Conservative: 3
 Best Local Similarity: 97.18% Mismatches: 1
 Query Match: 80.09% Indels: 1
 DB: 13 Gaps: 0

US-09-817-198B-2 (1-212) x BI557933 (1-699)

Qy 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyValGly 20
 Db 156 ATGCGGAACAGTACGATGCTGTCGGCTACTGCTGATCGGGAGCTCGGGGTGGC 215
 Qy 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40
 Db 216 AAGACATGCTGCTGTGCGCTTCACGACAACAGATTCCACTCTCGCATATCTCCACC 275
 Qy 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
 Db 276 ATCGGTGTTGACTTAAAGATGAAGACTATCGATAGACGGCATCAAAGTGAGATACAG 335
 Qy 61 IleThrAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
 Db 336 ATTTGGGACACAGCAGGCGAGGAGGTACCACTATCACAAGACAGTACTATCGGGCA 395
 Qy 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
 Db 396 GCCCAGGGAATATTTTAGTCTACGACATTAAGCAGTACGAGCGCTCTATCAGCATATCATG 455
 Qy 101 LysThrValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
 Db 456 AAGTGGGTCACTGAGTGGATGAGTACGCTCCAGAAAGAGTCCAGAAAGATCCTAATTTGGG 515
 Qy 121 AsnLysAlaAspGluGlnLysArgGlnVal-GlyArgGluGlnGlyGlnLeuAl 140
 Db 516 AATCAGGCTGATGAAGACAGACAACGGCAGCGTGGGCGACAGAGCGGGCAGCAGTGGC 575
 Qy 140 aLysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGI 160
 Db 576 TAAGAGTACGGCATGAGCTTCTACGAACAAGTCCCTGCAACCACTTATATATAAGA 635
 Qy 160 uSerPheThrArgLeuThrGluLeuValLeuAlaHisArgLysGlu 176
 Db 636 GTCTTCACTGCTGACGGAGTGGTCTGACGGCCACAGGACACAG 684

RESULT 7
 BF966292
 LOCUS 602286692f1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4375366 5',
 DEFINITION 796 bp mRNA linear EST 23-JAN-2001
 mRNA sequence.
 ACCSSION BF966292
 VERSION BF966292.1 GI:12333507
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 796)
 TITLE NIH-MGC http://mgi.nci.nih.gov/.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

FEATURES
 source


```

QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisLeuSerThr 40
Db 220 AAGACATGCGTGTGTGCGCTTCACGACGACAGAGTTCACCTCTCGCATATCTCCACC 279
QY 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
Db 280 ATCGGTGTTGACTTTAAAGATGAAGACTATGATGTAGACGGCATCAAGAGTGAAGATACAG 339
QY 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTrpArgArg 80
Db 340 ATTTGGACACACAGCGCAGGAGAGGTACACACTATCACAAAGCAGTACTATCGCGGA 399
QY 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
Db 400 GCCCAGGGAATATTTTAGTCTACGACATTAGCAGTGGCGCTCTATCAGCATATCATG 459
QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db 460 AAGTGGTCACTGACGTGGATGATGAGTACGCTCCAGAGGAGTCCAGAAAGATCCTAATTGGG 519
QY 121 AsnLysAlaAspGluGlnGlnLysArgGlnVal-GlyArgGluGlnGlyGlnGlnLeuAl 140
Db 520 AATAAGGCTGATCAAGACGACAAACGCGCAGTGGGGAGCAGACGAGGGGAGCAGCTGGC 579
QY 140 aLysGluTyrGlyMetaspPheTyrGluThrSerAlaCysThrAsnLeuAsn-IleLys 159
Db 580 TAAGGAGTACGGCAGTACTTCTACGAAACAAAGTGGCTGGAACAACTTAACTATTACAC 639
QY 160 GluSerPheThrArg-LeuThrGluLeuValLeuGlnAlaHisArgLys-GluLeuGluG 179
Db 640 GATCCTTCACTCGCTTGACGAGCTGGTCTGCGAGGCCACAGGAAGCAGCTGGACT 699
QY 179 lLeuArgMetArgAlaSerAsnGluLeu-----AlaLeuAlaGluL 193
Db 700 GG-----GCTCCGACACAGCGTGGCCACAGCAGCGCGCGCACTGGGCGAAC 747
QY 193 eu-----GluGluGluGlyLysProGluGly-----ProAlaAsnSers 207
Db 748 TTGGGAGGACGGACGACAGGAAACCCCGAGGGGGCCCGGACAAACCTTCCCAAGCA 807
QY 207 er 207
Db 808 CC 809

RESULT 9
BG247902
LOCUS
DEFINITION
602359819F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488197 5',
mRNA sequence.
ACCESSION
BG247902
VERSION
BG247902.1 GI:12757717
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 616)
AUTHORS
NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL
National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10334 row: f column: 06
High quality sequence stop: 616.
Location/Qualifiers

```

```

source
1. 616
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4488197"
/clone_lib="NCI_CGAP_Mam1"
/tissue_type="tumor, biopsy sample"
/dev_stage="10 months, virgin"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"
BASE COUNT 165 a 152 c 183 g 116 t
ORIGIN
Alignment Scores:
Pred. No.: 7.73e-96 Length: 616
Score: 829.00 Matches: 158
Percent Similarity: 99.38% Conservative: 3
Best Local Similarity: 97.53% Mismatches: 1
Query Match: 75.02% Indels: 0
DB: 12 Gaps: 0
US-09-817-198B-2 (1-212) x BG247902 (1-616)
QY 51 GluValAspGlyIleLysValArgIleGlnIleTrpAspThrAlaGlyGlnGluArgTyr 70
Db 3 GATGTAGACGGCATCAAGTGAGATACAGATTGGGACACAGCAGGCGAGGAGGTAC 62
QY 71 GlnThrIleThrLysGlnTyrTyrArgAlaGlnGlyIlePheLeuValTyrAspIle 90
Db 63 CAGACTATCAACAAGCAGTACTATCGCGAGCCCGAGGAATATTTTAGTCTACGACATT 122
QY 91 SerSerGluArgSerTyrGlnHisIleMetLysTrpValSerAspValAspGluTyrAla 110
Db 123 AGCAGTGGCGCTCTATCAGCATATCATGAAGTGGTCTAGTACGAGGATGATGAGTAC 182
QY 111 ProGluGlyValGlnLysIleLeuIleGlyAsnLysAlaAspGluGlnLysArgGln 130
Db 183 CCAGAGGAGTCCAGAAAGATCCTAATTGGGAATAAGCTGATGAAGACGAGAAACGCGAG 242
QY 131 ValGlyArgGluGlnGlyGlnGlnLeuAlaLysGluTyrGlyMetaspPheTyrGluThr 150
Db 243 GTGGGAGAGACGAGGCGGCGACAGCTGGCTGAAGAGTACGCGATGGACTTCTACGAAACA 302
QY 151 SerAlaCysThrAsnLeuAsnIleLysGluSerPheThrArgLeuThrGluLeuValLeu 170
Db 303 AGTGCCTGCGACCAACCTTAATATTAAGAGTCTCTACTCTGTCGCGGAGCTGGTCTG 362
QY 171 GlnAlaHisArgLysGluLeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeu 190
Db 363 CAGGCCACAGGAAAGAGCTGGATGGTCTCCGACACAGTCCACAGCAGAGCTCGACTG 422
QY 191 AlaGluLeuGluGluGluGluGlyLysProGluGlyProAlaAsnSerSerLysThrCys 210
Db 423 GCCGAGCTGGAGGAGGACGAAGGCAAACTTACGAGGCCCGAGCAAACTTCTCAAGACCTGC 482
QY 211 TrpCys 212
Db 483 TGGTGC 488

RESULT 10
BF182001
LOCUS
DEFINITION
601805936F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4036664 5',
mRNA sequence.
ACCESSION
BF182001
VERSION
BF182001.1 GI:11060143
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1121)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

[illegible][illegible]

Alignment Scores:	
Pred. No.:	1121
Score:	816.00
Percent Similarity:	85.92%
Best Local Similarity:	82.63%
Query Match:	73.85%
DB:	12
Length:	1121
Matches:	176
Conservative:	7
Mismatches:	25
Indels:	6
Gaps:	2

US-09-817-198B-2 (1-212) x BF182001 (1-1121)		FEATURES	source	Location/Qualifiers
QY	1	MetaAlaLysGlnTyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValcIcIy 20	1..814	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4860437" /clone_lib="NIH_MGC_49" /tissue_type="melanotic melanoma, high MDR (cell line)" /lab_host="DH10B (phage-resistant)" /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCACGAG(G). Size-selected >500bp for average insert size of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library. "
Db	196	ATGCCGAACAGTACGATGCTGCTTCGGCTACTGCTGATCGGGGACTCGGGGTGGC 255		
QY	21	LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40		
Db	256	AAGACATGCTGCTGCTGCCGCTTCCACGACACAGAGTCCACTCCTCGCATATCTCCACC 315		
QY	41	IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLeysValArgIleGln 60		
Db	316	ATCGGTGTGACTTTAAGATGAAGACTATCGAGTACGCGGCATCAAGTCAGATACAG 375		
QY	61	IleTrpAspThrAlaGlyGlnGlu-ArgTyrGlnThrIleThrLysGlnTyrTyrArgAr 80		
Db	376	ATTTTGGGACACAGCAGGGCAGCAGGATACCAGACTATCAAGCAGTAGTACTATCGGCG 435		

	80	gAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMe	100
Qy	80	gAlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMe	100
Db	436	AGCCCAAGGAATATTTTGTAGCTACGACATTAGCAGTGGCGCTCTATCAGCATATCAT	495
Qy	100	tlYsTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleG1	120
Db	496	GAAGTGGGTGAGTGCAGTGGATGAGTACGCTCCAGAAGAGTCCAGAAGATCCTAATNGG	555
Qy	120	yAsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGlnGlnGlnLeuAl	140
Db	556	GAATAAGGCTGATGATGAAGACGACAAACGCGAGTGGGGAGAGAGCAGGGCGCAGAG-CTGGCC	614
Qy	140	alYsGluTyrG1VmetAspPheTyrGluThr-SerAlaCysThrAsnLeuAsnIleLysG	160

ORIGIN

Alignment Scores:

pred. No.: 8.53e-92

Score: 799.00

Percent Similarity: 96.43%

Best Local Similarity: 95.83%

Query Match: 72.31%

DB: 12

US-09-817-198B-2 (1-212) x BG762967 (1-814)

Length: 814

Matches: 161

Conservative: 1

Mismatches: 3

Indels: 3

Gaps: 0

Qy

48 LysThrIleGluValAspGlyIleLysValArgIleGlnIleTrpAspThrAlaGlyGln

67

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAM11807 row: e column: 13

High quality sequence stop: 777.

FEATURES
source

1. 911
Location/Qualifiers
/organism="Mus musculus"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone_lib="IMAGE:5318388"
/clone_lib="NCI CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."

BASE COUNT 217 a 253 c 244 g 195 t 2 others
ORIGIN

Alignment Scores:

Pred. No.: 6.8e-89 Length: 911
Score: 777.00 Matches: 157
Percent Similarity: 98.16% Conservative: 3
Best Local Similarity: 98.32% Mismatches: 3
Query Match: 70.32% Indels: 2
DB: 13 Gaps: 0

US-09-817-198B-2 (1-212) x BF1649317 (1-911)

QY 50 lleGluValAspGlyileLysValArgIleGlnIleTTPAspThrAlaGlyGlnGluArg 69
|||||
Db 1 ATCGATGTAGACGGCATCAAGTGAGATACAGATTGGGACACAGCGCAG-GAGAGG 59
QY 70 TyrGlnThrIleThrLysGlnTyrTyrArgArgAlaGlnGlyIlePheLeuValTyrAsp 89
|||||
Db 60 TACCAGATATACAAAGCAGTACATCGGGAGGCCA-GGAATATTTTGTCTACGAC 118
QY 90 IleSerSerGluArgSerTyrGlnHisIleMetLysTrpValSerAspValAspGluTyr 109
|||||
Db 119 ATTACGATGACGGCTCTCTATCAGCATATCATGAAGTGGGTGAGTGGATGAGTAC 178
QY 110 AlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAlaAspGluGlnLysArg 129
|||||
Db 179 GCTCCAGAGGAGTCCAGAAATCCTAATTGGGAATAAGGCTGATGAAGAGCAGAAACGG 238
QY 130 GlnValGlyArgGluGlnGlyGlnGlnLeuAlaLysGluTyrGlyMetAspPheTyrGlu 149
|||||
Db 239 CAGGTGGGGAGAGACAGGGGAGCAGGTGGCTAAGGAGTACGGCATGGACTTCTACGAA 298
QY 150 ThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThrArgLeuThrGluLeuVal 169
|||||
Db 299 ACAAGTGGCTGCACCACTTAATTAATTAAGAGTCTCTCCTCGTCTCAGGAGTGGTG 358
QY 170 LeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAla 189
|||||
Db 359 CTGCAGGCCACAGAGAAAGAGTGGTGTCTCCGAACACAGCGCCAGCAACAGCTCGCA 418
QY 190 LeuAlaGluLeuGluGluGluGlyLysProGluGlyProAlaAsnSerSerLysThr 209
|||||
Db 419 CTGGCCGAGCTGGAGGAGGACGAGCAACCTGAGGGCCCGCCAGCAACTCTTCAAGACC 478
QY 210 CysTrpCys 212
|||||
Db 479 TCCTGGTGC 487

RESULT 14

BF150465

LOCUS

BF150465 618 bp mRNA linear EST 29-DEC-2000

DEFINITION

uy86b05.y1 NCI CGAP_Mam5 Mus musculus cDNA clone IMAGE:3666417 5' similar to SW:RB15_RAT P35289 RAS-RELATED PROTEIN RAB-15. ; mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT

EMAIL

TISSUE

CDNA

LIBRARY

ARRANGED

BY

THE

I.M.A.G.E.

CONSORTIUM

(LLNL)

CLONE

DISTRIBUTION

NCI CGAP

CLONE

DISTRIBUTION

INFORMATION

CAN

BE

FOUND

THROUGH

THE

I.M.A.G.E.

CONSORTIUM

(LLNL)

AT:

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

image.llnl.gov/image/html/iresources.shtml

|||||
397 GCCAGGGGAATATTAGTCTTAGGACATTAGCAGTGGCTCCCTATCAGCATATCATG 456
QY |||||
101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
Db |||||
457 AAGTGGGTGAGTACGCTGGATGAGTACGCTCCAGAGAGAGTCCAGAGATCCCTAATTGGG 516
QY |||||
121 AsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
Db |||||
517 AATAATGCTGATGAAGAGCAGAAACCGCAGGTGGGAGAGAGAGAGTGCAGCAGCCTGGC 576
QY |||||
141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThr 154
Db |||||
577 TTAGAGTACGCGATGGACTTCTAGGAACAGTGCCTGCACC 618

RESULT 15
BM462147
LOCUS BM462147 1767 bp mRNA linear EST 05-FEB-2002
DEFINITION ACENCOURT_6424626 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5491242
5', mRNA sequence.
ACCESSION BM462147
VERSION BM462147.1 GI:18511187
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo..
REFERENCE 1 (bases 1 to 1767)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cygabs-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12111 row: g column: 19
High quality sequence start: 89
High quality sequence stop: 574.
Location/Qualifiers
1. 1767
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5491242"
/clone_lib="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SF0RT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

BASE COUNT 318 a 704 c 411 g 320 t 14 others
ORIGIN

Alignment Scores:
Pred. No.: 8.62e-83 Length: 1767
Score: 733.00 Matches: 141
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 66.33% Indels: 0
DB: 13 Gaps: 0

US-09-817-198B-2 (1-212) x BM462147 (1-1767)

QY 72 ThrileThrLysGlnTyrTyrArgAlaGlnGlyPheLeuValTyrAspIleSer 91
Db |||||
96 ACCATCAAAAGCAGTACTATCGCGCGCCAGGGATATTTTGGTCTATGACATTAGC 155
QY 92 SerGluArgSerTyrGlnHisIleMetLysTrpValSerAspValAspGluTyrAlaPro 111

Db 156 AGCGAGCGCTTTACCAGCACATCATGAAGTGGTCACTGACGTGGATGATGACGACCA 215
QY |||||
112 GluGlyValGlnLysIleLeuIleGlyAsnLysAlaAspGluGluGlnLysArgGlnVal 131
Db |||||
216 GAAGCGCTCCAGAGATCCCTTATTTGGGAATAGCTCATGAGGAGCAGAACGCGAGTG 275
QY |||||
132 GlyArgGluGlnGlyGlnGlnLeuAlaLysGluTyrGlyMetAspPheTyrGluThrSer 151
Db |||||
276 GGAAGAGAGCAAGGCGCAGCAGCTGGCGAAGAGTATGGCATGGACTTCTATGAACAAGT 335
QY |||||
152 AlaCysThrAsnLeuAsnIleLysGluSerPheThrArgLeuThrGluLeuValLeuGln 171
Db |||||
336 GCCTGCACCAACCTCAACATTAAGAGTCAATTCACGCTCTGACAGAGCTGGTCTGCAG 395
QY |||||
172 AlaHisArgLysGluLeuGluGlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAla 191
Db |||||
396 GCCCATAGGAAGAGAGCTGGGAAGCGCTCCGGATGGCTGCCAGCATGATGGCTGGCA 455
QY |||||
192 GluLeuGluGluGluGluGlyLysProGluGlyProAlaAsnSerSerLysThrCysTrp 211
Db |||||
456 GAGCTGGAGGAGGAGGAGGAGGCAACCGGCGCCAGCAACTCTTCGAAAACCTGCTGG 515
QY 212 Cys 212
Db 516 TGC 518

Search completed: June 23, 2003, 12:15:52
Job time : 1390 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 23, 2003, 08:15:02 ; Search time 42 Seconds
(without alignments)
1547.987 Million cell updates/sec

Title: us-09-817-198b-2
Perfect score: 1105
Sequence: 1 MAKQDVLFRLILIGDVG.....LEEREGKPGPANSKTCWC 212

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n.model -DEV-xmlh
-Q-cgn2_1/USPTO_spool/US09817198/runat_18062003_144633_27928/app_query.fasta_1.391
-DB-Issued Patents_NA -QFMT-fastap -SUFFIX-p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09817198 -CGN_1_1_40 -runat_18062003_144633_27928 -NCPU=6 -ICPU=3
-NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents_NA:**

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:**
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:**
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:**
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:**
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:**
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	478.5	43.3	925	2	US-08-916-901-4
2	478.5	43.3	925	4	US-09-154-602-4
3	437.5	39.6	1340	2	US-08-824-873-2
4	437.5	39.6	1340	3	US-09-198-184-2
5	429	38.8	875	4	US-09-075-454-10
6	429	38.8	2612	4	US-09-484-970B-142
7	419	37.9	639	4	US-09-399-913-66
8	411.5	37.2	970	3	US-08-888-077A-28
9	373	33.8	847	2	US-08-773-423-4
10	343	31.0	890	3	US-08-741-411-4
11	340.5	30.8	803	4	US-09-075-454-13
12	321.5	29.1	1407	4	US-09-493-914-1
					Sequence 4, Appl1
					Sequence 4, Appl1
					Sequence 2, Appl1
					Sequence 2, Appl1
					Sequence 10, Appl1
					Sequence 142, App
					Sequence 66, Appl
					Sequence 28, Appl
					Sequence 4, Appl1
					Sequence 4, Appl1
					Sequence 13, Appl1
					Sequence 1, Appl1

13	320	29.0	1172	4	US-09-075-454-8	Sequence 8, Appl1
14	320	29.0	1533	4	US-09-075-454-11	Sequence 11, Appl1
15	313	28.3	1255	2	US-08-766-551-6	Sequence 6, Appl1
16	297	26.9	1175	2	US-08-773-423-6	Sequence 2, Appl1
17	282	25.5	848	3	US-08-741-411-2	Sequence 2, Appl1
18	274.5	24.8	1749	4	US-09-149-476-54	Sequence 54, Appl1
19	271.5	24.6	820	3	US-08-741-411-6	Sequence 6, Appl1
20	262	23.7	615	1	US-08-247-946A-5	Sequence 5, Appl1
21	262	23.7	615	5	PCT-US95-06420-5	Sequence 5, Appl1
22	260	23.5	985	4	US-08-842-306B-1	Sequence 1, Appl1
23	260	23.5	985	4	US-08-838-973B-1	Sequence 1, Appl1
24	260	23.5	985	4	US-08-771-212A-1	Sequence 1, Appl1
25	259.5	23.5	3198	4	US-08-842-306B-48	Sequence 48, Appl1
26	259.5	23.5	3198	4	US-08-838-973B-48	Sequence 48, Appl1
27	255	23.1	600	3	US-09-078-317-1	Sequence 1, Appl1
28	255	23.1	600	4	US-09-454-818-1	Sequence 1, Appl1
29	255	23.1	2309	3	US-09-078-317-3	Sequence 3, Appl1
30	255	23.1	2309	4	US-09-454-818-3	Sequence 3, Appl1
31	249.5	22.6	624	4	US-09-415-522-1	Sequence 3, Appl1
32	249.5	22.6	5197	4	US-09-293-170-6	Sequence 1, Appl1
33	246	22.3	1074	2	US-09-156-424-1	Sequence 1, Appl1
34	246	22.3	1074	4	US-09-387-341-1	Sequence 1, Appl1
35	245.5	22.2	603	4	US-09-325-932A-29	Sequence 29, Appl1
36	245.5	22.2	932	4	US-09-325-932A-28	Sequence 28, Appl1
37	245.5	22.2	2436	1	US-08-306-891B-16	Sequence 16, Appl1
38	244.5	22.1	574	2	US-08-429-964-83	Sequence 83, Appl1
39	239	21.6	607	2	US-08-429-964-85	Sequence 85, Appl1
40	236	21.4	1058	3	US-09-156-807-1	Sequence 1, Appl1
41	236	21.4	5775	1	US-08-306-691B-15	Sequence 15, Appl1
42	236	21.4	5775	5	PCT-US93-06251-29	Sequence 29, Appl1
43	232	21.0	1284	2	US-09-161-015-1	Sequence 1, Appl1
44	232	21.0	1284	4	US-09-387-341-150	Sequence 150, App
45	230.5	20.9	1166	5	PCT-US96-12129B-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1

US-08-916-901-4
; Sequence 4, Application US/08916901
; Patent No. 5892012
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: RAB PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,901
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0367 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166

QY 121 AsnLysAlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
 |||||:||||| |||||: |||
 Db 426 AACAGAGCGGCTTACCACCAAGAGGTGGTGGCAACACACACAGCAAGAGGTTTGA 485
 |||||:||||| |||||: |||
 QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
 |||||:||||| |||||: |||
 Db 486 GACTCTCTGGCATCCCTTCTTGGAGACGAGGCCCAAGATGCCACCAATGTGGACGAG 545
 |||||:||||| |||||: |||
 QY 161 SerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgTysGluLeuGluGlyLeu 180
 |||||:||||| |||||: |||
 Db 546 GCGTTC-----ATGACCATGCTCTGAAATCAAAAGCGGATGGGCGCT 590
 |||||:||||| |||||: |||
 QY 181 ArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluLeuGluGluGlyLysPro 200
 |||||:||||| |||||: |||
 Db 591 GGACGACGCTCTGGGGGCGAGCGGCCCAATCTCAAGATCGACACGCCCTGTAAAGCCG 650
 |||||:||||| |||||: |||
 QY 201 Glugly 202
 |||||: |||
 Db 651 GCTGGC 656
 |||||: |||

RESULT 3
 US-08-824-873-2
 ; Sequence 2, Application US/08824873
 ; Patent No. 5843717
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Guegler, Karl
 ; TITLE OF INVENTION: NOVEL RAB PROTEIN
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA: US/08/824,873
 ; FILING DATE: Filed Herewith
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0240 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1340 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: FANCMOT04
 ; CLONE: 738957
 US-08-824-873-2

Alignment Scores: 6.5e-46 Length: 1340
 Pred. No.: 90
 Score: 437.50 Matches: 90
 Percent Similarity: 67.84% Conservative: 45
 Best Local Similarity: 45.23% Mismatches: 52
 Query Match: 39.59% Indels: 13
 Db: 2 Gaps: 3

US-09-817-198B-2 (1-212) x US-08-824-873-2 (1-1340)
 QY 5 TyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyValGlyLysThrCysLeu 24
 |||||:||||| |||||: |||
 Db 22 TACAGCTGCGCTTCAAGGTCTGCTGGTGGGAGACTCGGTGTGGGAGAACCTGTCTG 81
 |||||:||||| |||||: |||
 QY 25 Leu-CysArgPheThrAspAsnGluPhe---HisSerSerHisIleSerThrIleGlyVa 43
 |||||:||||| |||||: |||
 Db 82 CTGGGTGCGATTCAAGGATGTGTCTCTCGCGGGGACCTTCATCTCCACCGTAGC-AT 140
 |||||:||||| |||||: |||
 QY 43 LaspPheLysMetLysThrIleGluValAlaGlyLysValGlyValArgIleGlnIleTrpAs 63
 |||||:||||| |||||: |||
 Db 141 TGACTTCGGGAACAAAGTCTGGAGCTGGATGGTGAAGTGAAGTGCAGATGTGGGA 200
 |||||:||||| |||||: |||
 QY 63 pThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgAlaGlnGln 83
 |||||:||||| |||||: |||
 Db 201 CACAGCTGGTCAGGCGGTTCCGAGTGTGTACCCATGCTACTACCGGGATGCTCATGC 260
 |||||:||||| |||||: |||
 QY 83 yIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMetLysTrpVa 103
 |||||:||||| |||||: |||
 Db 261 TCTGCTGCTGCTTACGATGTCAACACAGGCGCTCTCTTGCACACATCCAGGCGCTG 320
 |||||:||||| |||||: |||
 QY 103 lSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAl 123
 |||||:||||| |||||: |||
 Db 321 GACCGAGATCCAGGAGTACGCCAGCAGCGTGGCGCTCATGCTGTGGGGAACAAGGT 380
 |||||:||||| |||||: |||
 QY 123 aAspGluGlnGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAlaLysGluTy 143
 |||||:||||| |||||: |||
 Db 381 GGACTCTGCCCATGAGCGTGTGTGAGAGGAGGAGCGGGAGAGCTGGCCNAGAGTA 440
 |||||:||||| |||||: |||
 QY 143 rGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheTh 163
 |||||:||||| |||||: |||
 Db 441 TGGACTGCCCTTCATGGAGACCGCGCCCAAGCGGCGCTCAACGTGGAGCTTGGCCTT 500
 |||||:||||| |||||: |||
 QY 163 rArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetAr 183
 |||||:||||| |||||: |||
 Db 501 AGCCATAGCA-----AAGAGTGAAGCAGCGCTCCATGAA 536
 |||||:||||| |||||: |||
 QY 183 gAlaSerAsnGlu-----LeuAlaLeuAlaGluLeuGluGluGluGlyLys 199
 |||||:||||| |||||: |||
 Db 537 GGCTCCCGAGCGCGCGCTTCCGCTGCATGATTACGTTAAGAGGGAGGTCGA 591
 |||||:||||| |||||: |||

RESULT 4
 US-09-198-184-2
 ; Sequence 2, Application US/09198184
 ; Patent No. 6010859
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Guegler, Karl
 ; TITLE OF INVENTION: NOVEL RAB PROTEIN
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/198,184
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/824,873
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.


```
Db 312 TTGCTTCTGCTGATGACATCACCACAAATCTTTCGACAAACATCAGGGCTGGCTC 371
QY 104 SerAspValaspGluYrAlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAla 123
Db 372 ACTGAGATTTCATGATGATGCCAGAGGAGCGTGTGATCATGCTGCTAGGCAACAGGCG 431
QY 124 AspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLysGluTyr 143
Db 432 GATATGAGCAGCGAAGAGAGTGTGCTCCGAAAGACGAGAGACCTTGGCCAGGAGTAC 491
QY 144 GlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThr 163
Db 492 GGTGTTCCCTTCTCTGGAGACCGCCAGACCTGGCATGAATGTG----- 536
QY 164 ArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArg 183
Db 537 -----GAGTTAGCTTCTTGGCCATCGCCCAAGGAA-----CTGAATAACCGG 578
QY 184 AlaSerAsnGlu 187
Db 579 GCCGGGCATCAG 590
```

RESULT 6

```
US-09-484-970B-142
; Sequence 142, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 142
; LENGTH: 2612
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 412477.1CB1
US-09-484-970B-142
```

```
Alignment Scores:
Pred. No.: 2,08e-44 Length: 2612
Score: 429.00 Matches: 82
Percent Similarity: 68.48% Conservative: 44
Best Local Similarity: 44.57% Mismatches: 46
Query Match: 38.82% Indels: 12
DB: 4 Gaps: 3
```

US-09-817-198B-2 (1-212) x US-09-484-970B-142 (1-2612)

```
QY 5 TyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValGlyLysThrCysLeu 24
Db 91 TAGCGACTCAGCGCAAGGTGATGCTCTGGGAGACACAGCGCTCGCAAAACATGTTTC 150
QY 25 LeuCysArgPheThrAspAsnGluPheHisSer---SerHisIleSerThrIleGlyVal 43
Db 151 CTGATCCATTTAAAGACGGGGCTTCCGTGTCGGGAACCTTCATAGCCACCGTCGCGCAT 210
QY 44 AspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGlnIleTyrAsp 63
Db 211 GACTTCAGGACAAAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 270
QY 64 ThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrArgArgAlaGlnGly 83
Db 271 ACCGCTGGCAGGACGGTCCGAGAGCGTCACCCATGCTATTACAGAGATGCTCAGGCC 330
QY 84 IlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMetLysTrpVal 103
```

```
Db 331 TTGCTTCTGCTGATGACATCACCACAAATCTTTCGACAAACATCAGGGCTGGCTC 390
QY 104 SerAspValaspGluYrAlaProGluGlyValGlnLysIleLeuIleGlyAsnLysAla 123
Db 391 ACTGAGATTTCATGATGATGCCAGAGGAGCGTGTGATCATGCTGCTAGGCAACAGGCG 450
QY 124 AspGluGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLysGluTyr 143
Db 451 GATATGAGCAGCGAAGAGAGTGTGCTCCGAAAGACGAGAGACCTTGGCCAGGAGTAC 510
QY 144 GlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGluSerPheThr 163
Db 511 GGTGTTCCCTTCTCTGGAGACCGCCAGACCTGGCATGAATGTG----- 555
QY 164 ArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGluGlyLeuArgMetArg 183
Db 556 -----GAGTTAGCTTCTTGGCCATCGCCCAAGGAA-----CTGAATAACCGG 597
QY 184 AlaSerAsnGlu 187
Db 598 GCCGGGCATCAG 609
```

RESULT 7

```
US-09-399-913-66
; Sequence 66, Application US/09399913
; Patent No. 6361971
; GENERAL INFORMATION:
; APPLICANT: Rhodes, Kenneth
; APPLICANT: Betty, Maria
; APPLICANT: Ling, Hual-Ping
; APPLICANT: An, Wenqian
; TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR
; FILE REFERENCE: MNI-070CP2
; CURRENT APPLICATION NUMBER: US/09/399,913
; CURRENT FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER: USSN 60/110,277
; EARLIER FILING DATE: 1998-11-30
; EARLIER APPLICATION NUMBER: USSN 60/110,033
; EARLIER FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: USSN 60/109,333
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: USSN 09/298,731
; EARLIER FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: USSN 09/350,614
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: USSN 09/350,874
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 66
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Rattus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(636)
US-09-399-913-66
```

Alignment Scores:

```
Pred. No.: 4.8e-44 Length: 639
Score: 419.00 Matches: 85
Percent Similarity: 64.59% Conservative: 50
Best Local Similarity: 40.67% Mismatches: 65
Query Match: 37.92% Indels: 9
DB: 4 Gaps: 2
```

US-09-817-198B-2 (1-212) x US-09-399-913-66 (1-639)

```
QY 5 TyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValGlyLysThrCysLeu 24
Db 7 TAGCGCTAATCTCTTCAAGTACATCATCGGACACAGGTGTTGTTAAATCGTCTTA 66
```


142	utyrGlywetAspPheTyrGluThrSerAlaCysThsAsnLeuAsnIleLysGluSerPh	162
486	GTATGGGCGATTATTTCTGTAAACACAGTGCACCAAGATGTTCTTAACATAGTGGAGCGTGT	545
162	eThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGlu-	176
546	TCTGCACCTGCTCGAGAAAGTAAAGAGAGAAGCTGACAAAGATGACAGCAGCATCCATTAC	605

RESULT 12

RECORDED 2
US-09-493-914-1
: Sequence 1, Application US/09493914
: Patent No. 6448073
: GENERAL INFORMATION:
: APPLICANT: Jger, Dick
: APPLICANT: Stockert, Elizabeth
: APPLICANT: Jger, Elke
: APPLICANT: Gure, Ali
: APPLICANT: Scanlan, Matthew J.
: APPLICANT: Knuth, Alexander
: APPLICANT: Old, Lloyd J.
: APPLICANT: Chen, Yao-Tseng

Alignment Scores:	
Pred. No.:	3.5e-31
Score:	321.50
Percent Similarity:	44.19%
Best Local Similarity:	36.56%
Query Match:	29.10%
DB:	4
Length:	1407
Matches:	83
Conservative:	40
Mismatches:	66
Indels:	38
Gaps:	10

US-09-817-198B-2 (1-212) x US-09-493-914-1 (1-1407)

8	LeuPheArgLeuLeuIleGlyAspSerGlyValGlyLysThrCysLeuLeuCysArg	27
	::::::::::::::::::::	
Db	72 CTGTACAAGTTGCTGGTGATTGGCGACCTGGCGGTGGGAAGACAGTATCATCAAGCGC	131
QY	28 PheThrAspAsnGlupHeHisSerSerHisIleSerThrIleGlyValAspPhetysMet	47
	::::::::::::::::::::	
Db	132 TACGTGCACCAGAAGACTTCCTCTGCATACCGGCCACAAATCGCGTGCAGCTTCGCGCTC	191
QY	48 LysThrIleGluValasp--GlyIleLysValArgIleGlnIleIleTrpAspThrAlaGly	66
	::::::::::::::::::::	
Db	192 AAGTGTCTCCACTGGGACCCGGAGACGTGGTGGCGCTGCAGCTCGGGATATCGCAGET	251
QY	67 GlnGluArgTyrGlnThrIleThrLysGlnTyTyrArgArgAlaGlnGlyIlePheLeu	86
	::::::::::::::::::::	
Db	252 CAAGAAAGATTGGAACATGACGAGGGTCTATTACCGAAGCTATGGTGCANTTATT	311
QY	87 ValTyrAspIleSerSerGluArgSerTyrGlnHisIleMethLyTrpValSerAspVal	106
	::::::::::::::::::::	
Db	312 GTCTTCGATGTCACCGGCAGCCACATTTGAAGCAGTGGCAAAGTGGAAAATGATTGG	371
QY	107 AspGlu-----TyrAlaProGluGly-----ValGlnLysIleLeuIleGlyAsnLys	122
	::::::::::::::::::::	
Db	372 GACTCCAAGTTAAGCTCCCTTAATGGTCAACACCGGTTTCAGTGGTTTTGTTGGCCAACAA	431

123	AlaAspGluGluGlnLysArgGlnValGlyArgGluGlnGly-----GlnGlnLeu	139
	::::	
Ddb	432 TGTGAC-----CAGGGAAAGGATGCTCATGAACAATGGCGCTCAAGATGGACCAGTTC	485
	:	
Qy	140 AlaLysGluTyrGly---MetAppPheTyrGluThrSerAlaCysThrAsnLeuAsnIle	158
	:	
Ddb	486 TGCAAGGAGCACGGTTTCGTAGGATGGTTTTGAACAATCAGCAAAGGAAAATAATAACATT	545
	:	
Qy	159 LysGluSerPheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLysGluLeuGlu	178
	:	
Ddb	546 GATGAAGCTCCAGATGCCGTGGTGAACAACACTACT-----581	
	:	
Qy	179 GlyLeuArgMetArgAlaSerAsnGluLeuAlaLeuAlaGlu---LeuGluGluGluGlu	197
	:	
Ddb	582 -----GCNAATGATGTGCACCTAATGGAGTCTATTGAGCGCGGACGTC	623
	:	
Qy	198 GlyLysPro-----GluGlyProAla-AsnSerSe	207
	:	
Ddb	624 GTGAAGCCCCATCTCATCATCAACCAAGGTTGCCAGCTGCTCTGGCGTGCCAAATCCTTAG	683
	:	
Qy	207 rLys---ThrCystTrpCys	212
	:	
Ddb	684 TAGGCACCTTTGCTGGTGT	702
	:	

RESULT 13

US-09-075-454-8
Sequence 8, Application US/09075454
Patent No. 6391580
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Tang, Y. Tom
APPLICANT: Lal, Preeti
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
APPLICANT: Patterson, Chandra
APPLICANT: Batra, Sajeev
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: RAS PROTEINS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/075,454
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/766,551
FILING DATE: DECEMBER 12, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Corrone, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0168-1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1172 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 23, 2003, 11:10:12 ; Search time 158 Seconds
(without alignments)
1968.953 Million cell updates/sec

Title: US-09-817-198B-2

Perfect score: 1105

Sequence: 1 MAKQDVLFRLLIGDVG.....LEEKGKPGPANSSKTCWC 212

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09817198/runat_18062003_14634_27974/app_query.fasta_1.391
-DB=Published Applications_NA -QEXT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62
-TRANS=human40.cdi -LIST=45 -LOCALIGN=200 -THR_SCORE=pcr -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pcr -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09817198_ECGN_1_1_57_erunat_18062003_14634_27974
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1105	100.0	2021	9	US-09-764-868-88 Sequence 88, Appl
2	1105	100.0	3257	10	US-09-817-198A-1 Sequence 1, Appl1
3	832	75.3	566	9	US-09-764-868-507 Sequence 507, Appl
4	534.5	48.4	1274	10	US-09-925-302-91 Sequence 91, Appl

5	530.5	48.0	1537	10	US-09-925-300-631 Sequence 631, App
6	528.5	47.8	624	10	US-09-794-257-9 Sequence 9, Appl1
7	528.5	47.8	1161	10	US-09-794-257-7 Sequence 7, Appl1
8	528.5	47.8	2497	10	US-09-834-975-879 Sequence 879, App
9	528.5	47.8	2497	10	US-09-834-975-885 Sequence 885, App
10	528.5	47.8	2497	10	US-09-834-975-894 Sequence 894, App
11	528.5	47.8	2497	10	US-09-834-975-896 Sequence 896, App
12	519	47.0	651	9	US-09-938-842A-836 Sequence 836, App
13	478.5	43.3	925	10	US-09-967-736-4 Sequence 4, Appl1
14	478.5	43.3	939	9	US-10-102-806-48 Sequence 48, Appl1
15	477.5	43.2	881	10	US-09-770-445-529 Sequence 529, App
16	468	42.4	609	9	US-09-938-842A-832 Sequence 832, App
17	456.5	41.3	1069	9	US-10-037-270-646 Sequence 646, App
18	448	40.5	585	10	US-09-917-800A-1461 Sequence 1461, Ap
19	438	39.6	446	9	US-09-918-995-22915 Sequence 22915, A
20	429	38.8	875	12	US-10-051-986-10 Sequence 10, Appl
21	429	38.8	1106	9	US-10-037-270-959 Sequence 959, App
22	429	38.8	1116	10	US-09-794-257-13 Sequence 13, Appl
23	429	38.8	2674	10	US-09-817-199A-1 Sequence 1, Appl1
24	427	38.6	2623	9	US-09-764-868-71 Sequence 71, Appl
25	426	38.6	1316	9	US-09-764-868-493 Sequence 493, App
26	424	38.4	833	9	US-10-037-270-426 Sequence 426, App
27	420	38.0	429	9	US-09-918-995-5605 Sequence 5605, Ap
28	419	37.9	639	10	US-09-350-874-66 Sequence 66, Appl
29	416	37.6	576	10	US-09-794-257-15 Sequence 15, Appl
30	415.5	37.6	894	10	US-09-770-445-487 Sequence 487, App
31	412.5	37.3	538	10	US-09-924-035A-794 Sequence 794, App
32	411.5	37.2	654	9	US-09-938-842A-2113 Sequence 2113, Ap
33	410	37.1	1000	9	US-09-764-868-488 Sequence 488, App
34	406	36.7	896	10	US-09-770-445-478 Sequence 478, App
35	404.5	36.6	3936	10	US-09-919-172-49 Sequence 49, Appl
36	399	36.1	1127	9	US-09-764-868-487 Sequence 487, App
37	395	36.1	3124	10	US-09-925-302-340 Sequence 340, App
38	393	35.7	844	10	US-09-770-445-646 Sequence 646, App
39	391	35.4	1673	9	US-09-954-531-1359 Sequence 1359, Ap
40	388.5	35.2	4083	9	US-09-817-182-1 Sequence 1, Appl1
41	387.5	35.1	654	9	US-09-938-842A-78 Sequence 78, Appl
42	387	35.0	771	10	US-09-728-445-652 Sequence 652, App
43	384.5	34.8	1042	10	US-09-954-456-563 Sequence 563, App
44	382.5	34.6	676	10	US-09-770-149-374 Sequence 374, App
45	381	34.5	645	9	US-09-938-842A-774 Sequence 774, App

ALIGNMENTS

RESULT 1
US-09-764-868-88
; Sequence 88, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764, 868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 88
; LENGTH: 2021
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-88

Alignment Scores:
Pred. No.: 4.47e-132 Length: 2021
Score: 1105.00 Matches: 212
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-817-198B-2 (1-212) x US-09-764-868-88 (1-2021)

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 1161

; TYPE: DNA

; ORGANISM: homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (18)...(641)

; OTHER INFORMATION: n = A,T,C or G

US-09-794-257-7

Alignment Scores:

Pred. No.:	3,15e-58	Length:	1161
Score:	528.50	Matches:	95
Percent Similarity:	80.68%	Conservative:	47
Best Local Similarity:	53.98%	Mismatches:	33
Query Match:	47.83%	Indels:	1
DB:	10	Gaps:	1

US-09-817-198B-2 (1-212) x US-09-794-257-7 (1-1161)

```
QY 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyValGly 20
    |||||
Db 18 ATGCGAAGACGTACGATTATCTTCAAGCTCTGCTGATCGCGACTCGGGGTAGGC 77
    |||||
QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisIleSerThr 40
    |||||
Db 78 AAGACCTGCCTCTGTTCCGCTTCTCAGAGAGCGCCTTCAACACACCTTCATCTCCACC 137
    |||||
QY 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
    |||||
Db 138 ATCGGAATTGATTTAAATTAAGACATAGACTAGTGAAGAAATTAAGCTTCAG 197
    |||||
QY 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
    |||||
Db 198 ATATGGACACAGCGGTGAGAAAGATTCGGAACATACACGACGCGTACTACAGAGGA 257
    |||||
QY 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
    |||||
Db 258 GCATGGGCATTTATGCTGGTCTATGACATCAAAATGAAATTCCTTGTGCAATATTTAA 317
    |||||
QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
    |||||
Db 318 AATGGATCAGAACATTAAGAGCATGCTCTTCCGATGTCGAAAGAAATGATCCTGGGT 377
    |||||
QY 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGlnGlnGlnGlnLeuAla 140
    |||||
Db 378 AACAAATGTATATGAATGACAAAGACAAAGAGTGTCAAAGAAAGAGGGGAGAGTAGCA 437
    |||||
QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
    |||||
Db 438 ATTGACTATGGGATTAATTTCTGGAGACAGCGCAAAATCCAGTGCAAATGTAGAAGAG 497
    |||||
QY 161 Ser---PheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLys 175
    |||||
Db 498 GCATTTTACACTTCACGAGATATAATGACAAAACTCAACAGAAA 545
    |||||
```

RESULT 8

US-09-834-975-879

; Sequence 879, Application US/09834975

; Patent No. US20020110815A1

; GENERAL INFORMATION:

; APPLICANT: Lillie, James

; APPLICANT: Bolt, Andrew

; APPLICANT: Van Huffel, Christophe

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS

; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY

; FILE REFERENCE: MRI-016B

; CURRENT APPLICATION NUMBER: US/09/834,975

; PRIOR FILING DATE: 2001-04-13

; PRIOR FILING DATE: 2000-04-14

; NUMBER OF SEQ ID NOS: 1046

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 879

; LENGTH: 2497

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(2497)

; OTHER INFORMATION: n = A,T,C or G

US-09-834-975-879

Alignment Scores:

Pred. No.:	9,58e-58	Length:	2497
Score:	528.50	Matches:	95
Percent Similarity:	80.68%	Conservative:	47
Best Local Similarity:	53.98%	Mismatches:	33
Query Match:	47.83%	Indels:	1
DB:	10	Gaps:	1

US-09-817-198B-2 (1-212) x US-09-834-975-879 (1-2497)

```
QY 1 MetAlaLysGlnTyrAspValLeuPheArgLeuLeuIleGlyAspSerGlyValGly 20
    |||||
Db 92 ATGCGAAGACGTACGATTATCTTCAAGCTCTGCTGATCGCGACTCGGGGTAGGC 151
    |||||
QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerHisIleSerThr 40
    |||||
Db 152 AAGACCTGCCTCTGTTCCGCTTCTCAGAGAGCGCCTTCAACACACCTTCATCTCCACC 211
    |||||
QY 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyIleLysValArgIleGln 60
    |||||
Db 212 ATCGGAATTGATTTAAATTAAGACATAGACTAGTGAAGAAATTAAGCTTCAG 271
    |||||
QY 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
    |||||
Db 272 ATATGGACACAGCGGTGAGAAAGATTCGGAACATACACGACGCGTACTACAGAGGA 331
    |||||
QY 81 AlaGlnGlyIlePheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
    |||||
Db 332 GCATGGGCATTTATGCTGGTCTATGACATCAAAATGAAATTCCTTGTGCAATATTTAA 391
    |||||
QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
    |||||
Db 392 AATGGATCAGAACATTAAGAGCATGCTCTTCCGATGTCGAAAGAAATGATCCTGGGT 451
    |||||
QY 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGlnGlnGlnGlnLeuAla 140
    |||||
Db 452 AACAAATGTATATGAATGACAAAGACAAAGAGTGTCAAAGAAAGAGGGGAGAGTAGCA 511
    |||||
QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
    |||||
Db 512 ATTGACTATGGGATTAATTTCTGGAGACAGCGCAAAATCCAGTGCAAATGTAGAAGAG 571
    |||||
QY 161 Ser---PheThrArgLeuThrGluLeuValLeuGlnAlaHisArgLys 175
    |||||
Db 572 GCATTTTACACTTCACGAGATATAATGACAAAACTCAACAGAAA 619
    |||||
```

RESULT 9

US-09-834-975-885

; Sequence 885, Application US/09834975

; Patent No. US20020110815A1

; GENERAL INFORMATION:

; APPLICANT: Lillie, James

; APPLICANT: Bolt, Andrew

; APPLICANT: Van Huffel, Christophe

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS

; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY

; FILE REFERENCE: MRI-016B

; CURRENT APPLICATION NUMBER: US/09/834,975

; CURRENT FILING DATE: 2001-04-13


```

QY 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyLysValArgIleGln 60
DB 190 ATCGGGTGGACTTCAAGATCCGACCATCGAGCTGGATGCGCAAACTATCAAACTCAG 249
QY 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
DB 250 ATCTGGGACACACGGCGCCAGGAGCGTTCGGGACCATCTCCAGCTACTACCGGGG 309
QY 81 AlaGlnGlyLysPheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
DB 310 GCTCATGGCATCATCGTGGTGTATGAGCTCACTGACGAGCAATCTTACGCCAAGTGAAG 369
QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
DB 370 CAGTGGCTGCGAGGAGATTGACCTCTTCTGGAGCGGCAAGCAATCAATAGCTCTCTGGTGGC 429
QY 121 AsnLysAlaAspGluGlnGlnLysArgGlnValGlyArgGluGlnGlnLeuAla 140
DB 430 AACAGAGCGACCTCACCCACCAAGAGGTGGTGACACACACGAGGAGGAGTTGCA 489
QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
DB 490 GACTCTCTGGGCATCCCTCTTCTGGAGCGGCGCAAGCAATGCCACCAATGTCGAGCAG 549
QY 161 SerPheThrArgLeuThrGluLeuValLeuAlaHisArgLysGluLeuGluGlyLeu 180
DB 550 GCCTTC-----ATGACCATGGCTGCTGAATCAATCAAAAGCGGATGGGCGCT 594
QY 181 ArgMetArgAlaSerAsnGluLeuAlaLeuAlaGluGluGluGluGluGlyLysPro 200
DB 595 GGAGCAGCCCTCTGGGGGCGAGCGGCCCAATCTCAAGATCGACAGCACCCTCTGAAGCGG 654
QY 201 GluGly 202
DB 655 GCTGGC 660

```

RESULT 15

US-09-770-445-529
 ; Sequence 529, Application US/09770445
 ; Patent No. US2002023281A1

GENERAL INFORMATION:

```

; APPLICANT: Goralach, Jörn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krickler, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; TITLE OF INVENTION: thaliana
; FILE REFERENCE: 2023US (PARA-012PRV)
; CURRENT APPLICATION NUMBER: US/09770,445
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,472
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 529
; LENGTH: 881
; TYPE: DNA

```

; ORGANISM: Arabidopsis thaliana
 US-09-770-445-529

Alignment Scores:
 Pred. No.: 7,72e-52 Length: 881
 Score: 477.50 Matches: 96
 Percent Similarity: 65.57% Conservatives: 43
 Best Local Similarity: 45.28% Mismatches: 55
 Query Match: 43.21% Indels: 18
 DB: 10 Gaps: 4

US-09-817-198B-2 (1-212) x US-09-770-445-529 (1-881)

```

QY 1 MetalAlaGlnTyrAspValLeuPheArgLeuLeuLeuIleGlyAspSerGlyValGly 20
DB 60 ATGAATCCTGAGTACGACTATCTTTCAGCTCTGCTTATCGGGGATTCGCGGTAGGC 119
QY 21 LysThrCysLeuLeuCysArgPheThrAspAsnGluPheHisSerSerHisIleSerThr 40
DB 120 AAGCTTGTCTCTTTTGAGATTCTCTGATGATCTTATAGAAAAGTTACATTAGCAGCT 179
QY 41 IleGlyValAspPheLysMetLysThrIleGluValAspGlyLysValArgIleGln 60
DB 180 ATTGGAGTCGATTTAAATTAGGACTGTGGAACAAGATGCAAAACAATTAAGCTCAA 239
QY 61 IleTrpAspThrAlaGlyGlnGluArgTyrGlnThrIleThrLysGlnTyrTyrArgArg 80
DB 240 ATTTGGACACTGCTGCTCAAGAACGGTTCAGGACTATTACTAGCAGTTACTACCGTGG 299
QY 81 AlaGlnGlyLysPheLeuValTyrAspIleSerSerGluArgSerTyrGlnHisIleMet 100
DB 300 GCACATGGAATTAATATTGTCTACGATGTACAGATGACAGAAAGCTTCAATAATGTCAAG 359
QY 101 LysTrpValSerAspValAspGluTyrAlaProGluGlyValGlnLysIleLeuIleGly 120
DB 360 CAATGGTTGAGTGAATGATCGTTATGCTAGTGACAAATGTCAACAACCTCTTGTGGA 419
QY 121 AsnLysAlaAspGluGlnLysArgGlnValGlyArgGluGlnGlnGlnLeuAla 140
DB 420 AACAAAGTCTGATCTTACTGAAAACAGAGCCATTCCTTATGAAACTGCCAAGGCTTTTGC 479
QY 141 LysGluTyrGlyMetAspPheTyrGluThrSerAlaCysThrAsnLeuAsnIleLysGlu 160
DB 480 GATGAAATCGGGATTCCTTTTATGGAGACTAGTCAAAAGATGCTACAAACGTAGAACAG 539
QY 161 SerPheThrArgLeuThrGluLeuValLeuAlaHisArgLysGluLeuGluGlyLeu 180
DB 540 GCTTTC-----ATGGCAATGCTGCTCATCCATCAAGAG----- 572
QY 181 ArgMetArgAlaSer-AsnGluLeuAlaLeuAlaGluLeuGluGluGluGlyLysPr 200
DB 573 AGAATG---GCTAGCCCAACAGCTGGGAATAATGCAA-----GA 608
QY 200 oGluGlyProAlaAsnSerSerLysThrCysTrp 211
DB 609 CCACCCGACCGTCGAGATCAGAGGACGCGCTGTGG 642

```

Search completed: June 23, 2003, 12:54:35

Job time : 161 secs

THIS PAGE BLANK (USPTO)